



# Stable coupling of the Yee scheme with a linear current model

Filipe da Silva, Martin Campos Pinto, Bruno Després, Stephane Heuraux

## ► To cite this version:

Filipe da Silva, Martin Campos Pinto, Bruno Després, Stephane Heuraux. Stable coupling of the Yee scheme with a linear current model. *Journal of Computational Physics*, 2015, 295, pp.24-45. hal-01016883

**HAL Id: hal-01016883**

**<https://hal.science/hal-01016883>**

Submitted on 2 Jul 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

# Stable coupling of the Yee scheme with a linear current model

Filipe da Silva<sup>a</sup>, Martin Campos Pinto<sup>b,c</sup>, Bruno Després<sup>c,b,\*</sup>, Stéphane Heuraux<sup>d</sup>

<sup>a</sup>*Instituto de Plasmas e Fusão Nuclear, Instituto Superior Técnico, Universidade de Lisboa, 1049-001 Lisboa, Portugal*

<sup>b</sup>*CNRS, UMR 7598, Laboratoire Jacques-Louis Lions, F-75005, Paris, France*

<sup>c</sup>*Sorbonne Universités, UPMC Univ Paris 06, UMR 7598, Laboratoire Jacques-Louis Lions, F-75005, Paris, France*

<sup>d</sup>*Institut Jean Lamour, UMR 7198 CNRS-University Lorraine, Vandoeuvre, France*

---

## Abstract

This work analyzes the stability of the Yee scheme for non stationary Maxwell's equations coupled with a linear current model. We show that the usual procedure may yield unstable scheme for physical situations that correspond to strongly magnetized plasmas in X-mode (TE) polarization. We propose to use first order clustered discretization of the vectorial product that gives back a stable coupling. We validate the schemes on some test cases representative of direct numerical simulations of X-mode in a magnetic fusion plasma.

**Keywords:** Maxwell's equation, fusion plasma, stable scheme, FDTD, linear current model.

---

## 1. Introduction

The understanding of macroscopic phenomenons resulting from turbulent processes in magnetic plasmas requires huge numerical simulations. Such simulations should capture the statistical properties so that one can realize relevant averaging in time and space. This is particularly true to explain the macroscopic wave behavior which takes place on time scales that diagnostics cannot measure. An example which motivates our work is the recovery of the smooth envelope of a probing beam: to be able to compare with the theory, we need to simulate the non stationary Maxwell equations with a linear current model, see equations (1)-(2) below, in a domain with very fine non homogeneities (since the underlying plasma is turbulent) and over a simulation time equivalent to a half-million wave periods. This is required to study in details the role of the wavenumber spectrum on the beam widening [6, 24], and is illustrated with a snapshot of a numerical calculation of a wave in a plasma in Fig. 1. The Gaussian shape is recovered at coarse scale using some time averaging, but at the same time very fine scale details are visible. Since the energy of the wave is conserved during its propagation (we will prove this property for the model used in this work), it seems unavoidable to ask for similar energy preservation properties for the numerical methods in order to capture both the fine and coarse scales in the computations.

The previous example corresponding to beam broadening induced by wave propagation in turbulent plasma is illustrative of diagnostics and wave heating systems used in magnetized fusion plasmas. Hot topics for the ITER design concern the wavenumber resolution of the so-called

---

\*Corresponding author. LJLL, Université Pierre et Marie Curie, Boîte courrier 187, 75252 Paris Cedex 05, France.

Email addresses: tanatos@ipfn.ist.utl.pt (Filipe da Silva), campos@ann.jussieu.fr (Martin Campos Pinto), despres@ann.jussieu.fr (Bruno Després), stephane.heuraux@univ-lorraine.fr (Stéphane Heuraux)

Preprint submitted to J. Comp. Phys.

July 1, 2014

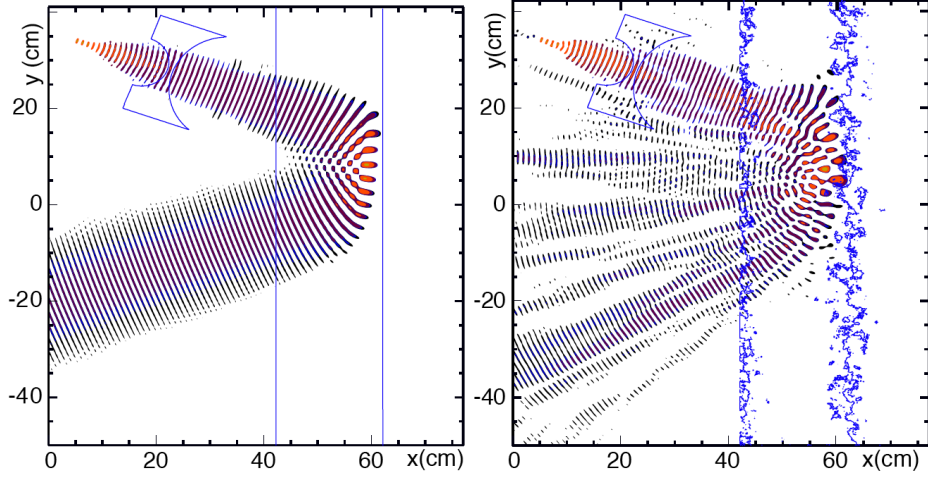


Figure 1: Contours at a given time of the non negative part of the electric field, that is  $\max(E_z, 0)$ , of an ordinary wave (known as TM or O-mode) with a Gaussian shape in unperturbed plasma with a linear density profile (left) and with turbulence of  $\delta n/n_c = 5\%$  (RMS value, with  $n_c$  the density at the cut-off layer for  $\nu = 40$  GHz) where the sub-beams are present (right).

Doppler reflectometry as well as the beam widening induced by the turbulence on the electron cyclotron heating to determine whether neo-classical tearing modes inducing big islands can be controlled or not [20]. The same questions arise for the lower-hybrid heating system to evaluate correctly what is the wavenumber spectrum launched into the plasma in view of the efficiency prediction of the non-inductive current driven system and the definition of the energy deposition zone. In the near future, wave polarization changes will be a subject of great interest, as it is becoming with O-X mode heating scenario for the stellarator or tokamak [17] and also for the new diagnostic concepts based on wave polarization changes induced by linear mode conversion or turbulence in inhomogeneous magnetized plasmas of space, astrophysics or fusion.

To our knowledge the state of the art of the computations done for such problems use simplified wave models taking into account only one physical mechanism (refraction in [20], diffraction effects in [27] and recently both in [15, 16]) but not all the wave structure as a Maxwell's equation solver with a linear current model can predict. This strongly questions the numerical stability of the simulations, indeed our numerical tests show a fundamental stability issue on extraordinary mode (also called TE or X-mode) which are more demanding in terms of stability than O-mode computations. Because an ideal scheme should be fast enough to reach the requirements stressed above for the comparison between theory and full-wave simulations, we are strongly interested in FDTD (Finite Difference Time Domain) schemes, indeed they are still the cheapests in terms of CPU. Many FDTD methods were developed to improve the performances and the stability in the past few years, such as EJ collocated FDTD [30], Runge-Kutta exponential time differencing formulation (RKETD) FDTD [19], matrix exponential (ME) FDTD [13], and exponential time differencing (ETD) FDTD [14]. Unfortunately they cannot be easily used to model the general dispersive media. On the other hand, in order to overcome the Courant-Friedrich-Levy (CFL) constraint of the conventional FDTD method, the one-step leapfrog ADI-FDTD method has been developed [4]. It originates from the conventional ADI-FDTD for which we refer to [31]. A re-

cent variant with higher computational efficiency is the implicit method [25]. We finally mention the fact that conservation of energy for a fourth-order FDTD scheme is addressed in [18], and that massive direct simulations of wave propagation in a different physical context can be found in the recent references [11, 9, 21].

The general model problem considered in this work is the non stationary Maxwell system

$$\begin{cases} -\varepsilon_0 \partial_t \mathbf{E} + \text{curl } \mathbf{H} = \mathbf{J} \\ \mu_0 \partial_t \mathbf{H} + \text{curl } \mathbf{E} = 0 \end{cases} \quad (1)$$

coupled with a linear equation for the electronic current density  $\mathbf{J} = eN_e \mathbf{u}_e$ ,

$$m_e \partial_t \mathbf{u}_e = e (\mathbf{E} + \mathbf{u}_e \wedge \mathbf{B}_0). \quad (2)$$

The unknowns are the electromagnetic field  $(\mathbf{E}, \mathbf{H})$ , with the usual notation that  $\mathbf{H} = \frac{\mathbf{B}}{\mu_0}$ , and the electronic current is  $\mathbf{J}$ . Equation (2) is the linearized Newton law for electrons where the Lorentz force on the right hand side is obtained by linearization [2, 3, 12, 22] of the full Newton law in a strongly magnetized medium with bulk magnetic field  $\mathbf{B}_0 \neq 0$ , and vanishing electric field and current  $\mathbf{E}_0 = \mathbf{J}_0 = 0$ . The coefficients are the mass  $m_e$  and charge  $e < 0$  of the electron, the permittivity  $\varepsilon_0$  and permeability  $\mu_0$  of vacuum. The electronic density  $N_e = N_e(t, \mathbf{x})$  is a prescribed function of the space and time variables. The background magnetic field could as well be dependent of the space and time variables  $\mathbf{B}_0 = \mathbf{B}_0(t, \mathbf{x})$ . The plasma frequency and cyclotron frequency are defined by

$$\omega_p = \sqrt{\frac{e^2 N_e}{m_e \varepsilon_0}} \quad \text{and} \quad \omega_c = \frac{|e \mathbf{B}_0|}{m_e}. \quad (3)$$

At any time, the total energy of the system is naturally the sum of the electromagnetic energy and the kinetic energy of electrons

$$\mathcal{E} = \int_{\Omega} \left( \frac{\varepsilon_0 |\mathbf{E}|^2}{2} + \frac{|\mathbf{B}|^2}{2\mu_0} + \frac{m_e |\mathbf{J}|^2}{2|e|N_e} \right) d\mathbf{x}. \quad (4)$$

This quantity will play a key role in our analysis.

The specific numerical difficulty addressed here is the stability of such couplings in the regime where the electronic density  $N_e$  has extremely strong gradients in space

$$|\nabla N_e| \gg 1. \quad (5)$$

The reason for such a strong gradient is that the underlying plasma is turbulent and strongly non homogeneous. See for example the density profiles used in the numerical section. On the other hand the gradient of  $\mathbf{B}_0$  can be considered small, even zero in first approximation. One may also consider that  $N_e$  is constant in time since this very simplified situation is representative of some reflectometry problems of fusion plasmas in X-mode polarization, see the simulations on a Cartesian grid and coupled with the Yee scheme [29] reported in [1, 7] and references therein. See also [28]. This hypothesis also simplifies the analysis since it can be checked that the total energy (4) is constant in time in this regime, up to the boundary contributions. It has been observed (see e.g. Fig. 6 below) that numerical instabilities sometimes develop in the regime  $\partial_t N_e = 0$  with long simulations using well-established coupling strategies between a Yee scheme discretizing

(1) and the linear current equation (2). For simplicity reasons and since it does not restrict the physics neither the numerics, the bulk of this article is devoted to the case where  $\partial_t N_e = 0$ , whereas the time dependent case is addressed in Appendix A. In the context of coupling the linear current equation (2) with Maxwell's equations (1), these numerical instabilities which are clearly related to strong gradients of the electronic density can also be characterized with the existence of different cut-offs and resonances which show up because the electronic density  $N_e$  varies with space, the source terms at the boundary are time harmonic with pulsation  $\omega$  and the time of simulation is large enough to reach the harmonic regime. Standard theory of waves in plasma [22, 23] pinpoints the role of the cyclotron resonance where  $\omega = \omega_c(\mathbf{x})$  and the hybrid-resonance where  $\omega^2 = \omega_p(\mathbf{x})^2 + \omega_c(\mathbf{x})^2$ . Since for a given  $\omega$  there may always exist a solution  $\mathbf{x}$  to these equations, it can yield a singular solution in time harmonic domain [10] and so be the germ of numerical instabilities for time domain simulations. This phenomenon will be evidenced in our numerical section for the hybrid resonance. The cut-off  $\omega = \omega_p(\mathbf{x})$  is not a singularity but a mild transition for propagative zone to elliptic (non propagative) and that is why it does not seem to yield the same kind of numerical difficulties as resonance do.

To be more specific let us assume now that  $\mathbf{b} = \frac{\mathbf{B}_0}{|\mathbf{B}_0|} = (0, 0, 1)^t$  is constant. It is not difficult to see that the initial problem (1)-(2) can be decomposed in two different polarizations [22, 23]. The first polarization is transverse magnetic so the electric field together with the electric current is always parallel to  $\mathbf{b}$ . In this case the vector product  $\mathbf{b} \wedge \mathbf{J}$  vanishes and the equation can be simplified a lot. This situation that does not yield resonances but only cut-off is called O-mode in plasma physics. The other polarization is transverse electric so that  $\mathbf{b} \wedge \mathbf{J}$  is a priori non zero. This polarization is called X-mode in the language of plasma physics. It is fundamentally the one we are interested in since the cyclotron resonance  $\omega = \omega_c(\mathbf{x})$  and the hybrid-resonance  $\omega^2 = \omega_p(\mathbf{x})^2 + \omega_c(\mathbf{x})^2$  can show up as the result of the resonant interaction of  $\mathbf{b} \wedge \mathbf{J} \neq 0$  with the rest of the system. The equations write

$$\begin{cases} -\varepsilon_0 \partial_t E_x + \partial_y H_z &= J_x, & J_x = e N_e u_x, \\ \varepsilon_0 \partial_t E_y - \partial_x H_z &= J_y, & J_y = e N_e u_y, \\ \mu_0 \partial_t H_z + \partial_x E_y - \partial_y E_x &= 0, \\ m_e \partial_t u_x &= e E_x + e u_y B_z^0, \\ m_e \partial_t u_y &= e E_y - e u_x B_z^0. \end{cases} \quad (6)$$

This system, as the previous one (1)-(2), is linear and will be used for the numerical tests, even if the analysis is done for the full 3D problem.

The main contributions of this work are in two directions: first we develop a theoretical framework that provides stable couplings in the energy norm between the Yee scheme (in any dimension) and a linear current model. Second, we design two simple explicit coupling schemes that are proved to be stable in the energy norm, a property that is generally not properly satisfied by existing FDTD schemes for such problems. The stability is achieved for all electronic densities, including cases with cut-offs or resonances in the computational domain. We will also prove that the CFL condition of one of our method is exactly the same as the one of the Yee scheme for Maxwell's equations in vacuum. Numerical tests performed for (6) with physically based coefficients will illustrate the properties of the new scheme.

The plan is as follows. In section 2 we recall the basics of the Yee scheme coupled with a linear current model. We shall outline the formal procedure that allows to derive explicit schemes and apply a standard stability analysis to different leap-frog time discretizations. Because in Yee schemes the different components of the vector fields are defined on staggered grids, care must be taken when defining the discrete curl and vector product operators. For this reason we

introduce abstract notations for these operators and we identify the conditions under which both the stability analysis and the formal derivation of an explicit scheme can be reproduced at the discrete level. In Section 3 we describe specific operators that satisfy these conditions, a property that is not fulfilled in usual FDTD schemes, see Remark 3.5. Equipped with these operators we then propose two explicit schemes that are proved to be stable under standard CFL conditions, and we give detailed formulas specifying the abstract operators. In Section 4 we finally show numerical experiments that demonstrate the improved stability of this approach.

## 2. Yee scheme with linear current model

The standard method to compute a numerical approximation of (1)-(2) on a Cartesian grid consists of coupling the Yee scheme with a linear discretization of the current, see [28, 1]. To have better insight in the structure of the coupled problem (1)-(2) we eliminate the velocity in terms of the current and discuss a discretization procedure. Here the data  $N_e$  is constant in time. Extension to the time-dependent case is postponed to Appendix A and Section 4.2.4.

Thus, the model problem is rewritten in a form which is common in plasma physics,

$$\begin{cases} -\varepsilon_0 \partial_t \mathbf{E} + \text{curl } \mathbf{H} = \mathbf{J}, \\ \mu_0 \partial_t \mathbf{H} + \text{curl } \mathbf{E} = 0, \\ \partial_t \mathbf{J} = \varepsilon_0 \omega_p^2(\mathbf{x}) \mathbf{E} + \omega_c \mathbf{b}(\mathbf{x}) \wedge \mathbf{J}. \end{cases} \quad (7)$$

As above, in inhomogeneous steady-state plasmas  $\omega_p(\mathbf{x}) = \sqrt{\frac{|e|^2 N_e(\mathbf{x})}{m \varepsilon_0}}$  and  $\omega_c(\mathbf{x}) = \frac{|e \mathbf{B}_0(\mathbf{x})|}{m_e}$  are the plasma and cyclotron frequencies, and  $\mathbf{b}(\mathbf{x}) = \frac{\mathbf{B}_0(\mathbf{x})}{|\mathbf{B}_0(\mathbf{x})|}$  is the direction of the background magnetic field.

Let  $n$  denote the time step ( $t_n = n \Delta t$ ). In the standard approach of [28, 1], the electrical field is discretized on the integer time steps  $n, n+1, \dots$ , while the other fields are on the half time steps  $n - \frac{1}{2}, n + \frac{1}{2}, \dots$ . The algorithmic structure of the standard method reads then

$$\begin{cases} \varepsilon_0 \frac{\mathbf{E}^{n+1} - \mathbf{E}^n}{\Delta t} = \text{curl } \mathbf{H}^{n+\frac{1}{2}} - \mathbf{J}^{n+\frac{1}{2}}, \\ \mu_0 \frac{\mathbf{H}^{n+\frac{1}{2}} - \mathbf{H}^{n-\frac{1}{2}}}{\Delta t} = -\text{curl } \mathbf{E}^n, \\ \frac{\mathbf{J}^{n+\frac{1}{2}} - \mathbf{J}^{n-\frac{1}{2}}}{\Delta t} = \varepsilon_0 \omega_p^2 \mathbf{E}^n + \omega_c \mathbf{b} \wedge \frac{\mathbf{J}^{n+\frac{1}{2}} + \mathbf{J}^{n-\frac{1}{2}}}{2}. \end{cases} \quad (8)$$

The equation for  $\mathbf{J}_h$  can be obtained as the limit for small friction of the one from [28], and is equal to the one from [1].

### 2.1. Formal derivation of an explicit scheme

Explicit versions of the above scheme can be derived by solving for  $\mathbf{J}^{n+\frac{1}{2}}$  the third equation, which has the structure

$$(I - \theta \mathbf{b} \wedge) \mathbf{J}^{n+\frac{1}{2}} = \mathbf{W}^n \quad (9)$$

with  $\theta = \frac{1}{2} \Delta t \omega_c$  and  $\mathbf{W}^n = \Delta t \varepsilon_0 \omega_p^2 \mathbf{E}^n + \mathbf{J}^{n-\frac{1}{2}} + \frac{1}{2} \Delta t \omega_c \mathbf{b} \wedge \mathbf{J}^{n-\frac{1}{2}}$ . Formally this is done by observing that since  $|\mathbf{b}| = 1$ , the identity

$$(\mathbf{b} \wedge)^3 = -(\mathbf{b} \wedge) \quad (10)$$

holds in the sense of operators acting on vector fields, i.e.,  $(\mathbf{b} \wedge)^3 \mathbf{V} = -(\mathbf{b} \wedge) \mathbf{V}$  for all  $\mathbf{V}$ . Using (10) one easily verifies

$$(I - \theta \mathbf{b} \wedge)^{-1} = I + \frac{\theta}{1 + \theta^2} (\mathbf{b} \wedge) + \frac{\theta^2}{1 + \theta^2} (\mathbf{b} \wedge)^2,$$

hence a formal solution for Equation (9) is given by

$$\mathbf{J}^{n+\frac{1}{2}} = \left( I + \frac{\theta}{1 + \theta^2} (\mathbf{b} \wedge) + \frac{\theta^2}{1 + \theta^2} (\mathbf{b} \wedge)^2 \right) \mathbf{W}^n \quad (11)$$

and this is the argument that is used in standard approaches to derive explicit solvers for (8), see also Remark 3.5 below.

In the Yee scheme it is necessary to specify how the curl and vector product operators act on discrete vector fields defined on staggered grid, and from the above discussion we already see that a discrete analogue of property (10) is likely to play a key role in the derivation of explicit schemes.

## 2.2. Discretizing on staggered grids

We consider a space discretization of (8) on a staggered mesh with spatial step  $h = \Delta x = \Delta y = \Delta z$ , which means that  $\mathbf{E}$ ,  $\mathbf{H}$  and  $\mathbf{J}$  are not all on the same grid. To continue the construction, we assume that  $\mathbf{J}$  is discretized on the edges of a Cartesian 3D cubic mesh. The center of cubic cells are denoted as  $\mathbf{x}_{i,j,k} := (ih, jh, kh)$ , for all triplets  $(i, j, k) \in D \subset \mathbb{Z}^3$  where  $D$  implicitly refers to the domain of computation.

With usual notations the discretization of the electric field is

$$\mathbf{E}_h = \left( E_x|_{i,j+\frac{1}{2},k+\frac{1}{2}}, E_y|_{i+\frac{1}{2},j,k+\frac{1}{2}}, E_z|_{i+\frac{1}{2},j+\frac{1}{2},k} \right)_{(i,j,k) \in D},$$

and similarly the discrete current density has the form

$$\mathbf{J}_h = \left( J_x|_{i,j+\frac{1}{2},k+\frac{1}{2}}, J_y|_{i+\frac{1}{2},j,k+\frac{1}{2}}, J_z|_{i+\frac{1}{2},j+\frac{1}{2},k} \right)_{(i,j,k) \in D}.$$

As for the magnetic field, it is discretized on the middle of faces,

$$\mathbf{H}_h = \left( H_x|_{i+\frac{1}{2},j,k}, H_y|_{i,j+\frac{1}{2},k}, H_z|_{i,j,k+\frac{1}{2}} \right)_{(i,j,k) \in D}.$$

It is convenient to introduce some notations for the spaces of vector fields defined on these staggered grids. Thus, the space of edgewise vector fields will be denoted

$$\mathcal{V}_e = \left\{ \mathbf{V}_h = \left( V_x|_{i,j+\frac{1}{2},k+\frac{1}{2}}, V_y|_{i+\frac{1}{2},j,k+\frac{1}{2}}, V_z|_{i+\frac{1}{2},j+\frac{1}{2},k} \right)_{(i,j,k) \in D} \right\} \quad (12)$$

so that  $\mathbf{E}_h, \mathbf{J}_h \in \mathcal{V}_e$ . The space of facewise quantities is

$$\mathcal{V}_f = \left\{ \mathbf{V}_h = \left( V_x|_{i+\frac{1}{2},j,k}, V_y|_{i,j+\frac{1}{2},k}, V_z|_{i,j,k+\frac{1}{2}} \right)_{(i,j,k) \in D} \right\} \quad (13)$$

so that  $\mathbf{H}_h \in \mathcal{V}_f$ . For later purposes we also introduce a notation for the space of cellwise vector fields,

$$\mathcal{V}_c = \left\{ \mathbf{V}_h = \left( V_x|_{i,j,k}, V_y|_{i,j,k}, V_z|_{i,j,k} \right)_{(i,j,k) \in D} \right\}. \quad (14)$$

The space discretization of (8) is easily described with these notations.

- The discretization of  $\text{curl } \mathbf{H}$  is realized as in the classical Yee scheme [29] by a matrix  $R$  acting from facewise quantities into edgewise quantities

$$R : \mathcal{V}_f \longrightarrow \mathcal{V}_e, \quad (15)$$

see also Equation (54) below.

- Again as in the classical Yee scheme, the transpose operator  $R^t : \mathcal{V}_e \longrightarrow \mathcal{V}_f$  will be used to compute  $\text{curl } \mathbf{E}$ .
- Assuming that the normalized background field  $\mathbf{x} \mapsto \mathbf{b}(\mathbf{x})$  is known everywhere, the space discretization of (8) requires a discrete operator

$$\mathbf{b} \wedge_h : \mathcal{V}_e \rightarrow \mathcal{V}_e$$

to represent the vector product by a unitary vector field  $\mathbf{b}$  within the space of edgewise vector fields. The exact definition and the compatibility properties of this operator will be detailed in Section 3.1.

- Finally we will need a notation for the multiplication of an edgewise vector by a scalar continuous function  $\mathbf{x} \mapsto u(\mathbf{x})$ . We let

$$S(u) : \mathcal{V}_e \rightarrow \mathcal{V}_e$$

denote the corresponding discrete operator, where  $S$  is a reminder of its scalar nature. Being essentially a diagonal operator,  $S(u)$  can be defined without loss of generality as

$$\begin{cases} (S(u)\mathbf{V}_h)_x|_{i,j+\frac{1}{2},k+\frac{1}{2}} = u^e|_{i,j+\frac{1}{2},k+\frac{1}{2}} V_x|_{i,j+\frac{1}{2},k+\frac{1}{2}} \\ (S(u)\mathbf{V}_h)_y|_{i+\frac{1}{2},j,k+\frac{1}{2}} = u^e|_{i+\frac{1}{2},j,k+\frac{1}{2}} V_y|_{i+\frac{1}{2},j,k+\frac{1}{2}} \\ (S(u)\mathbf{V}_h)_z|_{i+\frac{1}{2},j+\frac{1}{2},k} = u^e|_{i+\frac{1}{2},j+\frac{1}{2},k} V_z|_{i+\frac{1}{2},j+\frac{1}{2},k}, \end{cases} \quad (16)$$

where  $u^e$  should be seen as a discrete representation of  $u$  on the edges of the grid. Of course a natural choice would consist of simply taking

$$\begin{cases} u^e|_{i,j+\frac{1}{2},k+\frac{1}{2}} = u(\mathbf{x}_{i,j+\frac{1}{2},k+\frac{1}{2}}) \\ u^e|_{i+\frac{1}{2},j,k+\frac{1}{2}} = u(\mathbf{x}_{i+\frac{1}{2},j,k+\frac{1}{2}}) \\ u^e|_{i+\frac{1}{2},j+\frac{1}{2},k} = u(\mathbf{x}_{i+\frac{1}{2},j+\frac{1}{2},k}) \end{cases} \quad (17)$$

and in this case there would be no need for a new notation. However, it will be seen in Section 3.2 that this choice does not allow to easily derive explicit schemes that are stable, for the reason that the resulting  $S(u)$  does not commute with a proper definition of the discrete operator  $\mathbf{b} \wedge_h$ .

With these notations, the space discretization of (8) writes

$$\begin{cases} \varepsilon_0 \frac{\mathbf{E}_h^{n+1} - \mathbf{E}_h^n}{\Delta t} = R \mathbf{H}_h^{n+\frac{1}{2}} - \mathbf{J}_h^{n+\frac{1}{2}} \\ \mu_0 \frac{\mathbf{H}_h^{n+\frac{1}{2}} - \mathbf{H}_h^{n-\frac{1}{2}}}{\Delta t} = -R^t \mathbf{E}_h^n \\ \frac{\mathbf{J}_h^{n+\frac{1}{2}} - \mathbf{J}_h^{n-\frac{1}{2}}}{\Delta t} = \varepsilon_0 S(\omega_p^2) \mathbf{E}_h^n + S(\omega_c) \mathbf{b} \wedge_h \frac{\mathbf{J}_h^{n+\frac{1}{2}} + \mathbf{J}_h^{n-\frac{1}{2}}}{2}. \end{cases} \quad (18)$$



To analyze further the stability of the scheme it remains to study specific operators. Note that if the scalar fields (e.g.,  $\omega_p$ ) are constant in space, one has  $S(\omega_p^2) = \omega_p^2$ , but in general this is not the case. Let us finally emphasize the fact that for the moment (18) is only an abstract discrete scheme: in particular it has not even been discussed whether (18) is an explicit or an implicit method. As we will show later, an explicit version is available for a proper design of the discrete operators that mimics the algebra (9)-(11), see e.g. Section 3.4 below for detailed formulas.

### 2.3. Stability analysis

Our goal is here to analyze the stability of the abstract scheme (18), and obtain some conditions that the various operators should satisfy to obtain stability in the energy norm. To this end we symmetrize the continuous system and define

$$\widehat{\mathbf{E}} = \frac{\mathbf{E}}{c}, \quad \widehat{\mathbf{J}} = \frac{\mathbf{J}}{\omega_p c \varepsilon_0} \quad \text{and} \quad \widehat{\mathbf{B}} = \mathbf{B} = \mu_0 \mathbf{H}.$$

Still in the case where the plasma is stationary, that is  $\partial_t \omega_p = 0$ , one gets

$$\begin{cases} \partial_t \widehat{\mathbf{E}} - c \operatorname{curl} \widehat{\mathbf{B}} = -\omega_p \widehat{\mathbf{J}} \\ \partial_t \widehat{\mathbf{B}} + c \operatorname{curl} \widehat{\mathbf{E}} = 0 \\ \partial_t \widehat{\mathbf{J}} = \omega_p \widehat{\mathbf{E}} + \omega_c \mathbf{b} \wedge \widehat{\mathbf{J}}. \end{cases} \quad (19)$$

It can be written under the symmetric form of a Friedrichs system

$$\partial_t U + A \partial_x U + B \partial_y U + C \partial_z U = MU \quad (20)$$

with unknown  $U = (\widehat{\mathbf{E}}, \widehat{\mathbf{B}}, \widehat{\mathbf{J}})^t$ . The blockwise structure of the matrices is

$$A = c \begin{pmatrix} 0 & \tilde{A} & 0 \\ \tilde{A}^t & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad B = c \begin{pmatrix} 0 & \tilde{B} & 0 \\ \tilde{B}^t & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad C = c \begin{pmatrix} 0 & \tilde{C} & 0 \\ \tilde{C}^t & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

with

$$\tilde{A} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & I \\ 0 & -I & 0 \end{pmatrix} \quad \tilde{B} = \begin{pmatrix} 0 & 0 & -I \\ 0 & 0 & 0 \\ I & 0 & 0 \end{pmatrix} \quad \tilde{C} = \begin{pmatrix} 0 & I & 0 \\ -I & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

and finally

$$M = \begin{pmatrix} 0 & 0 & -\omega_p I \\ 0 & 0 & 0 \\ \omega_p I & 0 & \tilde{M} \end{pmatrix} \quad \text{with} \quad \tilde{M} = \omega_c \begin{pmatrix} 0 & -b_z I & b_y I \\ b_z I & 0 & -b_x I \\ -b_y I & b_x I & 0 \end{pmatrix}.$$

In these notations the size of the identity matrix is defined in accordance with the size of the global matrix. The total physical energy in a domain  $\Omega$  is

$$\mathcal{E}(t) = \int_{\Omega} \frac{1}{2\mu_0} |U(t, \mathbf{x})|^2 d\mathbf{x} = \int_{\Omega} \left( \frac{\varepsilon_0 |\mathbf{E}(t, \mathbf{x})|^2}{2} + \frac{|\mathbf{B}(t, \mathbf{x})|^2}{2\mu_0} + \frac{m_e |\mathbf{J}(t, \mathbf{x})|^2}{2|e|N_e(\mathbf{x})} \right) d\mathbf{x}. \quad (21)$$

Taking the scalar product of (20) with  $U$ , and performing basic integration by parts in the computational domain  $\Omega$ , one gets the fundamental law that

$$\frac{d}{dt}\mathcal{E}(t) = 0 \quad (22)$$

which expresses the conservation of the physical energy, up to boundary conditions of course.

We now extend this analysis to the scheme (18). Here the discrete symmetrized variables

$$\widehat{\mathbf{E}}_h^n = \frac{\mathbf{E}_h^n}{c}, \quad \widehat{\mathbf{J}}_h^{n+\frac{1}{2}} = S\left(\frac{1}{\omega_p}\right) \frac{\mathbf{J}_h^{n+\frac{1}{2}}}{c\mathcal{E}_0} \quad \text{and} \quad \widehat{\mathbf{B}}_h^{n+\frac{1}{2}} = \mathbf{B}_h^{n+\frac{1}{2}} = \mu_0 \mathbf{H}_h^{n+\frac{1}{2}} \quad (23)$$

are solution of

$$\begin{cases} \frac{\widehat{\mathbf{E}}_h^{n+1} - \widehat{\mathbf{E}}_h^n}{\Delta t} = cR\widehat{\mathbf{B}}_h^{n+\frac{1}{2}} - S(\omega_p)\widehat{\mathbf{J}}_h^{n+\frac{1}{2}} \\ \frac{\widehat{\mathbf{B}}_h^{n+\frac{1}{2}} - \widehat{\mathbf{B}}_h^{n-\frac{1}{2}}}{\Delta t} = -cR'\widehat{\mathbf{E}}_h^n \\ \frac{\widehat{\mathbf{J}}_h^{n+\frac{1}{2}} - \widehat{\mathbf{J}}_h^{n-\frac{1}{2}}}{\Delta t} = S(\omega_p)\widehat{\mathbf{E}}_h^n + S(\omega_c)\mathbf{b} \wedge_h \frac{\widehat{\mathbf{J}}_h^{n+\frac{1}{2}} + \widehat{\mathbf{J}}_h^{n-\frac{1}{2}}}{2}. \end{cases} \quad (24)$$

We also consider the natural hypothesis that the discrete cross product satisfies

$$\langle \mathbf{V}_h, \mathbf{b} \wedge_h \mathbf{V}_h \rangle_h = 0, \quad \mathbf{V}_h \in \mathcal{V}_e \quad (25)$$

where

$$\langle \mathbf{V}_h, \mathbf{W}_h \rangle_h = h^3 \sum_{(i,j,k) \in D} \left( (V_x W_x)|_{i,j+\frac{1}{2},k+\frac{1}{2}} + (V_y W_y)|_{i+\frac{1}{2},j,k+\frac{1}{2}} + (V_z W_z)|_{i+\frac{1}{2},j+\frac{1}{2},k} \right)$$

is the natural  $L^2$  scalar product in  $\mathcal{V}_e$ , and similar definition with a shift of indices holds concerning the norms in  $\mathcal{V}_f$  and in  $\mathcal{V}_c$ .

This of course will depend on the precise definition of the operator  $\mathbf{b} \wedge_h$ , but we note that it is obviously true at the continuous level. We define the discrete pseudo-energy by

$$\tilde{\mathcal{E}}_h^n = \|\widehat{\mathbf{E}}_h^n\|_h^2 + \|\widehat{\mathbf{B}}_h^{n-\frac{1}{2}}\|_h^2 + \|\widehat{\mathbf{J}}_h^{n-\frac{1}{2}}\|_h^2 - \Delta t \langle \widehat{\mathbf{E}}_h^n, cR\widehat{\mathbf{B}}_h^{n-\frac{1}{2}} - S(\omega_p)\widehat{\mathbf{J}}_h^{n-\frac{1}{2}} \rangle_h. \quad (26)$$

Here  $\|\mathbf{V}_h\|_h = \sqrt{\langle \mathbf{V}_h, \mathbf{V}_h \rangle_h}$  is the natural  $L^2$  norm in  $\mathcal{V}_e$ , and later we will denote by  $\|R\|_h = \max_{\|\mathbf{V}_h\|_h \leq 1} \|R\mathbf{V}_h\|_h$  the associated matrix norm.

**Proposition 2.1.** *Assume the orthogonality (25). One has  $\tilde{\mathcal{E}}_h^{n+1} = \tilde{\mathcal{E}}_h^n$ .*

*Proof.* The proof, which is a matter of easy algebra, nevertheless helps to understand the role of the key property (25). We take the scalar product of the first (resp. second, resp. third) equation of (24) by  $\widehat{\mathbf{E}}_h^{n+1} + \widehat{\mathbf{E}}_h^n$  (resp.  $\widehat{\mathbf{B}}_h^{n+\frac{1}{2}} + \widehat{\mathbf{B}}_h^{n-\frac{1}{2}}$ , resp.  $\widehat{\mathbf{J}}_h^{n+\frac{1}{2}} + \widehat{\mathbf{J}}_h^{n-\frac{1}{2}}$ ). Then the difference

$$\Delta := \frac{1}{\Delta t} (\|\widehat{\mathbf{E}}_h^{n+1}\|_h^2 + \|\widehat{\mathbf{B}}_h^{n+\frac{1}{2}}\|_h^2 + \|\widehat{\mathbf{J}}_h^{n+\frac{1}{2}}\|_h^2 - \|\widehat{\mathbf{E}}_h^n\|_h^2 - \|\widehat{\mathbf{B}}_h^{n-\frac{1}{2}}\|_h^2 - \|\widehat{\mathbf{J}}_h^{n-\frac{1}{2}}\|_h^2)$$

reads

$$\begin{aligned}\Delta = & \langle cR\widehat{\mathbf{B}}_h^{n+\frac{1}{2}} - S(\omega_p)\widehat{\mathbf{J}}_h^{n+\frac{1}{2}}, \widehat{\mathbf{E}}_h^{n+1} + \widehat{\mathbf{E}}_h^n \rangle_h + \langle -cR'\widehat{\mathbf{E}}_h^n, \widehat{\mathbf{B}}_h^{n+\frac{1}{2}} + \widehat{\mathbf{B}}_h^{n-\frac{1}{2}} \rangle_h \\ & + \langle S(\omega_p)\widehat{\mathbf{E}}_h^n + \frac{\omega_c}{2}\mathbf{b} \wedge_h (\widehat{\mathbf{J}}_h^{n+\frac{1}{2}} + \widehat{\mathbf{J}}_h^{n-\frac{1}{2}}), \widehat{\mathbf{J}}_h^{n+\frac{1}{2}} + \widehat{\mathbf{J}}_h^{n-\frac{1}{2}} \rangle_h.\end{aligned}$$

Many terms on the right hand side cancel, in particular the last one due to property (25). One gets then

$$\begin{aligned}\Delta = & \langle R\widehat{\mathbf{B}}_h^{n+\frac{1}{2}} - S(\omega_p)\widehat{\mathbf{J}}_h^{n+\frac{1}{2}}, \widehat{\mathbf{E}}_h^{n+1} \rangle_h - \langle cR'\widehat{\mathbf{E}}_h^n, \widehat{\mathbf{B}}_h^{n-\frac{1}{2}} \rangle_h + \langle S(\omega_p)\widehat{\mathbf{E}}_h^n, \widehat{\mathbf{J}}_h^{n-\frac{1}{2}} \rangle_h \\ = & \langle cR\widehat{\mathbf{B}}_h^{n+\frac{1}{2}} - S(\omega_p)\widehat{\mathbf{J}}_h^{n+\frac{1}{2}}, \widehat{\mathbf{E}}_h^{n+1} \rangle_h - \langle cR\widehat{\mathbf{B}}_h^{n-\frac{1}{2}} - S(\omega_p)\widehat{\mathbf{J}}_h^{n-\frac{1}{2}}, \widehat{\mathbf{E}}_h^n \rangle_h\end{aligned}$$

which ends the proof of the claim.  $\square$

**Remark 2.2.** In the case  $\mathbf{J} \equiv 0$ ,  $\omega_p = \omega_c \equiv 0$ , this relation is exactly the one used to analyze the stability of the Yee scheme. Indeed one can show that  $\|\mathbf{R}\|_h \leq 4/h$  in 3D. So if the CFL condition  $c\Delta t < h/2$  is satisfied, then one gets that  $\tilde{\mathcal{E}}_h^n$  actually defines a norm on  $(\widehat{\mathbf{E}}_h^n, \widehat{\mathbf{B}}_h^{n-1/2})$ , which establishes the stability of the Yee scheme. It is in fact possible to get a better CFL condition. Assuming an initial condition on the discrete free divergence of  $\widehat{\mathbf{E}}_h^0$  and  $\widehat{\mathbf{B}}_h^{1/2}$ , one can prove [26, Theorem 9, p. 160] that

$$\|\mathbf{R}\|_h \leq 2\sqrt{3}/h \quad \text{in 3D.} \quad (27)$$

In this case the CFL condition becomes  $c\Delta t < h/\sqrt{3}$  which is less stringent. The seminal Yee contribution [29] points out  $c\Delta t < h\sqrt{3}$  which is only a necessary condition.

To continue the analysis of the general scheme we set  $\|\omega_p\|_{L^\infty} = \sup_{\mathbf{x}} |\omega_p(\mathbf{x})|$ . We observe that a Cauchy-Schwarz inequality yields

$$\begin{aligned}\tilde{\mathcal{E}}_h^n \geq & \|\widehat{\mathbf{E}}_h^n\|_h^2 + \|\widehat{\mathbf{B}}_h^{n-\frac{1}{2}}\|_h^2 + \|\widehat{\mathbf{J}}_h^{n-\frac{1}{2}}\|_h^2 - c\Delta t \|\mathbf{R}\|_h \|\widehat{\mathbf{E}}_h^n\|_h \|\widehat{\mathbf{B}}_h^{n-\frac{1}{2}}\|_h \\ & - \Delta t \|\omega_p\|_{L^\infty} \|\widehat{\mathbf{E}}_h^n\|_h \|\widehat{\mathbf{J}}_h^{n-\frac{1}{2}}\|_h,\end{aligned}$$

that is

$$\tilde{\mathcal{E}}_h^n \geq P_{\alpha,\beta} \left( \|\widehat{\mathbf{E}}_h^n\|_h, \|\widehat{\mathbf{B}}_h^{n-\frac{1}{2}}\|_h, \|\widehat{\mathbf{J}}_h^{n-\frac{1}{2}}\|_h \right), \quad \alpha = c\Delta t \|\mathbf{R}\|_h, \quad \beta = \Delta t \|\omega_p\|_{L^\infty}.$$

If the polynomial  $P_{\alpha,\beta}(u, v, w) = u^2 + v^2 + w^2 - \alpha uv - \beta uw$  is positive for all non zero values of  $(u, v, w)$ , then  $\tilde{\mathcal{E}}_h^n$  is (the square of) a norm. This is if and only if. Assuming for example that  $u \neq 0$ , one can define  $x = \frac{v}{u}$  and  $y = \frac{w}{u}$  so that  $P_{\alpha,\beta}(u, v, w) = x^2 + y^2 - \alpha x - \beta y + 1$ . A further reduction is based on the decomposition  $(x, y) = \gamma \vec{a} + \delta \vec{b}$  where two vectors are orthonormal:  $\vec{a} = \frac{1}{\sqrt{\alpha^2 + \beta^2}}(\alpha, \beta)$  and  $\vec{b} = \frac{1}{\sqrt{\alpha^2 + \beta^2}}(\beta, -\alpha)$ . Plugging this decomposition in the polynomial, one gets  $P_{\alpha,\beta}(u, v, w) = (\gamma^2 - \gamma \sqrt{\alpha^2 + \beta^2} + 1) + \delta^2$ . So  $P$  is positive for all non zero values of  $(u, v, w)$  if and only if the discriminant of the first term in parenthesis is negative, that is  $\alpha^2 + \beta^2 < 4$ .

Thus we obtain that if the CFL condition

$$\frac{\Delta t}{2} \sqrt{c^2 \|\mathbf{R}\|_h^2 + \|\omega_p\|_{L^\infty}^2} < 1 \quad (28)$$

is satisfied, then  $\tilde{\mathcal{E}}_h^n$  is actually (the square of) a norm, which establishes the numerical stability of the scheme (18).

Using the estimate (27) for the norm of the operator  $R$ , one gets an estimate on the CFL condition

$$\frac{\Delta t}{2} \left( \frac{12c^2}{h^2} + \|\omega_p\|_{L^\infty}^2 \right)^{\frac{1}{2}} < 1.$$

This formula shows that the actual CFL condition is influenced by local extrema of the plasma density  $N_e$ : the higher the plasma frequency  $\|\omega_p\|_{L^\infty}$ , the smaller the time step. On the other hand the cyclotron frequency does not show up in this CFL condition.

#### 2.4. Optimization of the CFL condition

Since the pseudo-energy (26) contains a term coupling the electric current with the electric field, one gets at the end of the analysis a term proportional to the maximal norm of the plasma frequency  $\|\omega_p\|_{L^\infty}$  in the CFL condition (28). In view of this CFL condition, it is natural to ask if a slightly different method could get rid of that term  $\|\omega_p\|_{L^\infty}$ , so that the CFL condition is equal to the one in vacuum,

$$\frac{\Delta t}{2} c \|R\|_h < 1. \quad (29)$$

The answer to this question is positive as is now shown. We consider the scheme

$$\begin{cases} \frac{\widehat{\mathbf{E}}_h^{n+1} - \widehat{\mathbf{E}}_h^n}{\Delta t} = cR\widehat{\mathbf{B}}_h^{n+\frac{1}{2}} - S(\omega_p) \frac{\widehat{\mathbf{J}}_h^{n+1} + \widehat{\mathbf{J}}_h^n}{2} \\ \frac{\widehat{\mathbf{B}}_h^{n+\frac{1}{2}} - \widehat{\mathbf{B}}_h^{n-\frac{1}{2}}}{\Delta t} = -cR'\widehat{\mathbf{E}}_h^n \\ \frac{\widehat{\mathbf{J}}_h^{n+1} - \widehat{\mathbf{J}}_h^n}{\Delta t} = S(\omega_p) \frac{\widehat{\mathbf{E}}_h^{n+1} + \widehat{\mathbf{E}}_h^n}{2} + S(\omega_c)\mathbf{b} \wedge_h \frac{\widehat{\mathbf{J}}_h^{n+1} + \widehat{\mathbf{J}}_h^n}{2} \end{cases} \quad (30)$$

and the pseudo-energy

$$\mathcal{E}_h^n = \|\widehat{\mathbf{E}}_h^n\|_h^2 + \|\widehat{\mathbf{B}}_h^{n-\frac{1}{2}}\|_h^2 + \|\widehat{\mathbf{J}}_h^n\|_h^2 - \Delta t \langle \widehat{\mathbf{E}}_h^n, cR\widehat{\mathbf{B}}_h^{n-\frac{1}{2}} \rangle_h. \quad (31)$$

**Proposition 2.3.** *The pseudo-energy (31) is preserved time steps after time steps by (30), hence the latter is endowed with the CFL condition in vacuum.*

*Proof.* The proof is immediate by taking the scalar product of the first equation of this system (resp. second, resp. third) by  $\widehat{\mathbf{E}}_h^{n+1} + \widehat{\mathbf{E}}_h^n$  (resp.  $\widehat{\mathbf{B}}_h^{n+\frac{1}{2}} + \widehat{\mathbf{B}}_h^{n-\frac{1}{2}}$ , resp.  $\widehat{\mathbf{J}}_h^{n+1} + \widehat{\mathbf{J}}_h^n$ ).  $\square$

**Remark 2.4.** *This CFL condition is independent of all plasma parameters, which may have a practical interest if  $\|\omega_p\|_{L^\infty}$  is very large.*

In the original variables the scheme (30) reads

$$\begin{cases} \varepsilon_0 \frac{\mathbf{E}_h^{n+1} - \mathbf{E}_h^n}{\Delta t} = R\mathbf{H}_h^{n+\frac{1}{2}} - \frac{\mathbf{J}_h^{n+1} + \mathbf{J}_h^n}{2} \\ \mu_0 \frac{\mathbf{H}_h^{n+\frac{3}{2}} - \mathbf{H}_h^{n+\frac{1}{2}}}{\Delta t} = -R'\mathbf{E}_h^{n+1} \\ \frac{\mathbf{J}_h^{n+1} - \mathbf{J}_h^n}{\Delta t} = \varepsilon_0 S(\omega_p^2) \frac{\mathbf{E}_h^{n+1} + \mathbf{E}_h^n}{2} + S(\omega_c)\mathbf{b} \wedge_h \left( \frac{\mathbf{J}_h^{n+1} + \mathbf{J}_h^n}{2} \right). \end{cases} \quad (32)$$

**Remark 2.5.** *The extension of (32) to the time dependent case  $\partial_t N_e \neq 0$  is addressed in Appendix A where a stable scheme is provided, and in Section 4.2.4 where preliminary numerical experiments are discussed.*

### 3. Staggered grid operators and explicit schemes

In this section we will design a vector product operator acting on edgewise vector fields

$$\mathbf{b} \wedge_h : \mathcal{V}_e \rightarrow \mathcal{V}_e$$

that allows to derive stable and explicit schemes when used in conjunction with a proper operator

$$S(u) : \mathcal{V}_e \rightarrow \mathcal{V}_e$$

for the multiplication of an edgewise vector field by a scalar field  $u$ .

#### 3.1. Stable vector products on staggered grids

In the previous sections we have identified two key properties that the discrete vector product operator  $\mathbf{b} \wedge_h$  should satisfy. The first one is linked with the energy stability of the scheme: it is (25). The second one is linked with the existence of a local procedure to solve Equation (9), and hence with the existence of an explicit version of the scheme: it is the analogue of (10) and reads

$$(\mathbf{b} \wedge_h)^3 \mathbf{V} = -(\mathbf{b} \wedge_h) \mathbf{V} \quad \text{for all } \mathbf{V}_h \in \mathcal{V}_e. \quad (33)$$

To construct discrete operators on the staggered grids it will be convenient to denote by  $\tau_x^\alpha$ ,  $\alpha \in \{-1, +1\}$ , the half cell translation operator in the  $x$  direction,

$$(\tau_x^\alpha v)|_{i',j',k'} := v|_{i'+\frac{\alpha}{2},j',k'} \quad \text{for } (i',j',k') \in \frac{1}{2}\mathbb{Z}. \quad (34)$$

Similarly we denote by  $\tau_y^\beta$  and  $\tau_z^\gamma$  the translations in the other directions. We then observe that any triplet  $\alpha, \beta, \gamma \in \{-1, +1\}$  defines a translation operator

$$\tau = \tau^{\alpha,\beta,\gamma} : \mathcal{V}_e \rightarrow \mathcal{V}_e \quad (35)$$

that maps any  $\mathbf{V}_h \in \mathcal{V}_e$  into  $\mathcal{V}_e$ ,

$$(\tau V)|_{i,j,k} := (\tau_y^\beta \tau_z^\gamma V_x, \tau_x^\alpha \tau_z^\gamma V_y, \tau_x^\alpha \tau_y^\beta V_z)|_{i,j,k} \quad \text{for } i, j, k \in \mathbb{Z} \quad (36)$$

see Figure 2. Of course these notations must be technically adapted near the boundaries of the domain of computation.

##### 3.1.1. Averaged (centered, second order) vector products

Since the components of  $\mathbf{V}_h \in \mathcal{V}_e$  are defined on the edges of the grid, it is natural to design  $\mathbf{b} \wedge_h$  through an averaging procedure. Specifically, one could define

$$\begin{cases} (\mathbf{b} \wedge_h \mathbf{V}_h)_x|_{i,j+\frac{1}{2},k+\frac{1}{2}} := \frac{1}{4} \left( \sum_{\alpha,\beta=\pm 1} b_y \tau_x^\alpha \tau_z^\beta V_z - \sum_{\alpha,\beta=\pm 1} b_z \tau_x^\alpha \tau_y^\beta V_y \right) \Big|_{i,j+\frac{1}{2},k+\frac{1}{2}} \\ (\mathbf{b} \wedge_h \mathbf{V}_h)_y|_{i+\frac{1}{2},j,k+\frac{1}{2}} := \frac{1}{4} \left( \sum_{\alpha,\beta=\pm 1} b_z \tau_y^\alpha \tau_x^\beta V_x - \sum_{\alpha,\beta=\pm 1} b_x \tau_y^\alpha \tau_z^\beta V_z \right) \Big|_{i+\frac{1}{2},j,k+\frac{1}{2}} \\ (\mathbf{b} \wedge_h \mathbf{V}_h)_z|_{i+\frac{1}{2},j+\frac{1}{2},k} := \frac{1}{4} \left( \sum_{\alpha,\beta=\pm 1} b_x \tau_y^\alpha \tau_z^\beta V_y - \sum_{\alpha,\beta=\pm 1} b_y \tau_x^\alpha \tau_z^\beta V_x \right) \Big|_{i+\frac{1}{2},j+\frac{1}{2},k} \end{cases} \quad (37)$$

and indeed such formulas are often encountered in the practical implementation of Yee schemes coupled with a linear current model, see Remark 3.5

**Proposition 3.1.** *The operator (37) is formerly second order. It satisfies (25) if  $\mathbf{b}$  is a constant vector; and if  $D = \mathbb{Z}^3$ , but not (33).*

*Proof.* The formulas (37) are centered so second order. The fact that (33) is not true is a standard consequence of the averaging technique. Indeed such an averaging enlarges the stencil, so at the end the stencil of  $(\mathbf{b} \wedge_h)^3$  is strictly greater than the one of  $\mathbf{b} \wedge_h$ , which implies the non satisfaction of (33). Finally property (25) is elementary to check on an infinite Cartesian grid.  $\square$

The fact that the averaged vector product does not satisfy relation (33) has an important consequence on the stability of the numerical scheme, see Remark 3.5 below.

### 3.1.2. Clustered (non-centered, first order) vector products

In order to satisfy (33), the main idea is to design an operator  $\mathbf{b} \wedge_h$  that is block diagonal, so that iterated applications on a given  $\mathbf{V}_h$  with a single non-zero entry should not result in vector fields with increasing number of non-zero entries. This is easily obtained by using local *clusters* of nodes and can be formalized as follows. Although the vector field  $\mathbf{b}$  is in general known everywhere, we first introduce a discrete representation of it on the cell centers,

$$\mathbf{b}^c = (b_x^c|_{i,j,k}, b_y^c|_{i,j,k}, b_z^c|_{i,j,k})_{(i,j,k) \in D} \in \mathcal{V}_c$$

that may or may not coincide with the values  $\mathbf{b}(\mathbf{x}_{i,j,k})$ , but in every case satisfies

$$|(\mathbf{b}^c|_{i,j,k})| = ((b_x^c|_{i,j,k})^2 + (b_y^c|_{i,j,k})^2 + (b_z^c|_{i,j,k})^2)^{\frac{1}{2}} = 1 \quad \text{for all } (i, j, k) \in D.$$

We next define the operator  $\mathbf{b} \wedge_h$  as  $\tau^{-1} \mathbf{b}^c \wedge \tau$ , i.e.,

$$\tau(\mathbf{b} \wedge_h \mathbf{V}_h) := \mathbf{b}^c \wedge (\tau \mathbf{V}_h) \quad \text{for all } \mathbf{V}_h \in \mathcal{V}_e \quad (38)$$

where we observe that the standard product “ $\wedge$ ” can be used since all three components of the vector fields in  $\mathcal{V}_c$  are defined on the nodes  $(i, j, k) \in D$ . Specifically, expanding the above definition gives

$$\begin{cases} (\mathbf{b} \wedge_h \mathbf{V}_h)_x|_{i,j+\frac{\beta}{2},k+\frac{\gamma}{2}} := b_y^c|_{i,j,k} V_z|_{i+\frac{\alpha}{2},j+\frac{\beta}{2},k} - b_z^c|_{i,j,k} V_y|_{i+\frac{\alpha}{2},j,k+\frac{\gamma}{2}}, \\ (\mathbf{b} \wedge_h \mathbf{V}_h)_y|_{i+\frac{\alpha}{2},j,k+\frac{\gamma}{2}} := b_z^c|_{i,j,k} V_x|_{i,j+\frac{\beta}{2},k+\frac{\gamma}{2}} - b_x^c|_{i,j,k} V_z|_{i+\frac{\alpha}{2},j+\frac{\beta}{2},k}, \\ (\mathbf{b} \wedge_h \mathbf{V}_h)_z|_{i+\frac{\alpha}{2},j+\frac{\beta}{2},k} := b_x^c|_{i,j,k} V_y|_{i+\frac{\alpha}{2},j,k+\frac{\gamma}{2}} - b_y^c|_{i,j,k} V_x|_{i,j+\frac{\beta}{2},k+\frac{\gamma}{2}}. \end{cases} \quad (39)$$

Clearly the resulting operator  $\mathbf{b} \wedge_h$  is block diagonal (the size of each block being  $3 \times 3$  as in the standard vector product), and hence the locality of the operator is guaranteed. The degrees of freedom which are coupled by this strategy are depicted with black circles on Figure 2.

**Proposition 3.2.** *The operator  $\mathbf{b} \wedge_h$  defined in (38) is a first order approximation of the exact operator. It satisfies both (25) and (33) for general  $\mathbf{b}$  and  $D \subset \mathbb{Z}^3$ .*

*Proof.* The proof of these various properties is an immediate consequence of the definition of the operator. First one has that

$$\langle \mathbf{b} \wedge_h \mathbf{V}_h, \mathbf{V}_h \rangle_h = \langle \tau(\mathbf{b} \wedge_h \mathbf{V}_h), \tau \mathbf{V}_h \rangle_h = \langle \mathbf{b}^c \wedge \tau \mathbf{V}_h, \tau \mathbf{V}_h \rangle_h = 0$$

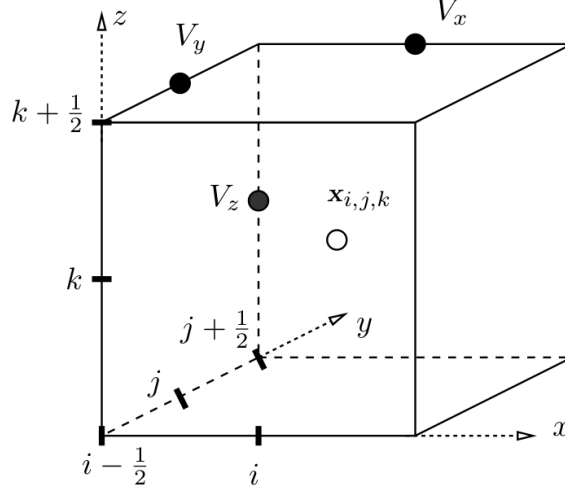


Figure 2: Graphical depiction of a cluster of staggered degrees of freedom defined by a translation  $\tau = \tau^{\alpha,\beta,\gamma}$  which pattern is  $(\alpha, \beta, \gamma) = (-1, +1, +1)$ . Here the vector field  $\tau \mathbf{V}_h$  is defined by (36) on each cell center  $\mathbf{x}_{i,j,k}$  (in white) from the components of  $\mathbf{V}_h$  located on the neighbouring edges (in black), namely  $V_x|_{i,j+\frac{1}{2},k+\frac{1}{2}}$ ,  $V_y|_{i-\frac{1}{2},j,k+\frac{1}{2}}$  and  $V_z|_{i-\frac{1}{2},j+\frac{1}{2},k}$ . From (38) we verify that the discrete operator  $\mathbf{b}_{\wedge h}$  is block-diagonal, with  $3 \times 3$  blocks corresponding to these clusters of components.

holds for any vector field  $\mathbf{b}$  and domain  $D$ , since every term involved in the  $(i, j, k)$ -sum underlying  $\langle \mathbf{b}^c \wedge \tau \mathbf{V}_h, \tau \mathbf{V}_h \rangle_h$  is zero. Next from  $|\mathbf{b}^c| = 1$  one verifies that  $(\mathbf{b}^c \wedge)^3 = -(\mathbf{b}^c \wedge)$ . Applying several times (38) this gives

$$\tau((\mathbf{b}_{\wedge h})^3 \mathbf{V}_h) = (\mathbf{b}^c \wedge)^3 \tau \mathbf{V}_h = -\mathbf{b}^c \wedge \tau \mathbf{V}_h = -\tau(\mathbf{b}_{\wedge h} \mathbf{V}_h)$$

which ends the formal proof of (33). The fact that the operator is first order is immediate.  $\square$

In particular, we observe that the formally less accurate definition (38) has stronger algebraic properties than the second order approximation (37).

### 3.2. Commutative multiplication by a scalar field

If  $\mathbf{x} \mapsto u(\mathbf{x})$  is a continuous scalar function, the most natural way to multiply some  $\mathbf{V}_h \in \mathcal{V}_e$  by  $u$  consists of applying the operator  $S(u)$  defined by (16) together with the choice (17). In general, however, the resulting  $S(u)$  does not commute with  $\mathbf{b}_{\wedge h}$ , which makes the resolution procedure (9)-(11) ineffective. Therefore, it seems natural to use a commutativity preserving operator, i.e., satisfying

$$S(u)(\mathbf{b}_{\wedge h} \mathbf{V}_h) = \mathbf{b}_{\wedge h} (S(u) \mathbf{V}_h) \quad \text{for all } \mathbf{V}_h \in \mathcal{V}_e. \quad (40)$$

To this end we easily verify the following result.

**Proposition 3.3.** *The operator  $S(u)$  defined by (16) satisfies (40) if and only if the edgewise representation  $u^e$  of the scalar field  $u$  satisfies*

$$u^e|_{i,j+\frac{\ell}{2},k+\frac{\gamma}{2}} = u^e|_{i+\frac{\alpha}{2},j,k+\frac{\gamma}{2}} = u^e|_{i+\frac{\alpha}{2},j+\frac{\ell}{2},k} \quad \text{for } (i, j, k) \in D, \quad (41)$$

where  $(\alpha, \beta, \gamma)$  is the triplet defining the translation operator  $\tau$  in (35) and the associated vector product  $\mathbf{b} \wedge_h$  in (38).

### 3.3. Explicit resolution

In this section we summarize the main results of this paper in a concise form using the staggered grid operators defined above. In particular we provide explicit versions of the resulting stable schemes, and refer to Section 3.4 below for detailed formulas involving the clustered (non-centered) operators introduced in Sections 3.1 and 3.2.

**Proposition 3.4.** *If the operators  $\mathbf{b} \wedge_h$  and  $S(u)$  satisfy (25), (33) and (40), then the schemes (18) and (32) are stable under the respective CFL conditions (28) and (29). Moreover, the scheme (18) has an explicit form given by*

$$\begin{cases} \mathbf{H}_h^{n+\frac{1}{2}} = \mathbf{H}_h^{n-\frac{1}{2}} - \frac{\Delta t}{\mu_0} R^t \mathbf{E}_h^n \\ \mathbf{J}_h^{n+\frac{1}{2}} = \left( I + S\left(\frac{\theta}{1+\theta^2}\right)(\mathbf{b} \wedge_h) + S\left(\frac{\theta^2}{1+\theta^2}\right)(\mathbf{b} \wedge_h)^2 \right) \mathbf{W}_h^n \\ \mathbf{E}_h^{n+1} = \mathbf{E}_h^n + \frac{\Delta t}{\varepsilon_0} (R \mathbf{H}_h^{n+\frac{1}{2}} - \mathbf{J}_h^{n+\frac{1}{2}}) \end{cases} \quad (42)$$

with  $\theta = \frac{\Delta t \omega_c}{2}$  and  $\mathbf{W}_h^n = \Delta t \varepsilon_0 S(\omega_p^2) \mathbf{E}_h^n + \left( I + \frac{\Delta t}{2} S(\omega_c) \mathbf{b} \wedge_h \right) \mathbf{J}_h^{n-\frac{1}{2}}$ . Finally the scheme (32) has an explicit form given by

$$\begin{cases} \mathbf{H}_h^{n+\frac{1}{2}} = \mathbf{H}_h^{n-\frac{1}{2}} - \frac{\Delta t}{\mu_0} R^t \mathbf{E}_h^n \\ \mathbf{J}_h^{n+1} = \left( I + S\left(\frac{\theta}{1+\theta^2}\right)(\mathbf{b} \wedge_h) + S\left(\frac{\theta^2}{1+\theta^2}\right)(\mathbf{b} \wedge_h)^2 \right) \mathbf{W}_h^n \\ \mathbf{E}_h^{n+1} = \mathbf{E}_h^n + \frac{\Delta t}{\varepsilon_0} (R \mathbf{H}_h^{n+\frac{1}{2}} - \frac{1}{2}(\mathbf{J}_h^{n+1} + \mathbf{J}_h^n)) \end{cases} \quad (43)$$

now with  $\theta = \frac{\Delta t \omega_c}{2\eta}$  where  $\eta = 1 + \left(\frac{\Delta t \omega_p}{2}\right)^2$ , and

$$\mathbf{W}_h^n = S\left(\frac{1}{\eta}\right) \left( \Delta t \varepsilon_0 S(\omega_p^2) (\mathbf{E}_h^n + \frac{\Delta t}{2\varepsilon_0} R \mathbf{H}_h^{n+\frac{1}{2}}) + \left( I - \frac{\Delta t^2}{4} S(\omega_p^2) + \frac{\Delta t}{2} S(\omega_c) \mathbf{b} \wedge_h \right) \mathbf{J}_h^n \right).$$

*Proof.* The stability properties have been established in Sections 2.3 and 2.4 under the assumption (25). To prove the explicit formulas we observe that (33) and (40) allow to reproduce the formal resolution procedure (9)-(11) at the discrete level. Indeed, using these relations it is easily verified that

$$\left( I + S\left(\frac{\theta}{1+\theta^2}\right)(\mathbf{b} \wedge_h) + S\left(\frac{\theta^2}{1+\theta^2}\right)(\mathbf{b} \wedge_h)^2 \right) (I - S(\theta) \mathbf{b} \wedge_h) \mathbf{V}_h = \mathbf{V}_h \quad (44)$$

holds for all  $\mathbf{V}_h \in \mathcal{V}_e$ . Thus, rewriting the last equation from (18) as

$$(I - S(\theta) \mathbf{b} \wedge_h) \mathbf{J}_h^{n+\frac{1}{2}} = \mathbf{W}_h^n \quad (45)$$

leads to the announced form (42). As for the new scheme (32), we observe that its implicit part reads

$$\begin{cases} \mathbf{E}_h^{n+1} = -\frac{\Delta t}{2\varepsilon_0} \mathbf{J}_h^{n+1} + \mathbf{U}_h^n, \\ \mathbf{J}_h^{n+1} = \frac{\Delta t \varepsilon_0}{2} S(\omega_p^2) \mathbf{E}_h^{n+1} + \frac{\Delta t}{2} S(\omega_c) \mathbf{b} \wedge_h \mathbf{J}_h^{n+1} + \mathbf{V}_h^n, \end{cases} \quad (46)$$



with source terms

$$\begin{cases} \mathbf{U}_h^n = \mathbf{E}_h^n + \frac{\Delta t}{\varepsilon_0} \mathbf{R} \mathbf{H}_h^{n+\frac{1}{2}} - \frac{\Delta t}{2\varepsilon_0} \mathbf{J}_h^n \\ \mathbf{V}_h^n = \mathbf{J}_h^n + \frac{\Delta t \varepsilon_0}{2} S(\omega_p^2) \mathbf{E}_h^n + \frac{\Delta t}{2} S(\omega_c) \mathbf{b} \wedge_h \mathbf{J}_h^n. \end{cases} \quad (47)$$

Eliminating  $\mathbf{E}_h^{n+1}$  in the second equation of (46) and applying  $S(\frac{1}{\eta})$ , we find

$$(I - \frac{\Delta t}{2} S(\frac{\omega_c}{\eta}) \mathbf{b} \wedge_h) \mathbf{J}_h^{n+1} = S(\frac{1}{\eta}) (\frac{\Delta t \varepsilon_0}{2} S(\omega_p^2) \mathbf{U}_h^n + \mathbf{V}_h^n) = \mathbf{W}_h^n \quad (48)$$

so that the desired form (43) is again obtained by applying the identity (44).  $\square$

**Remark 3.5.** *The fact that the averaged vector product (37) does not satisfy relation (33) – and in fact, is not block-diagonal – has two practical consequences: First, it means that if one replaces the formal vector products by (37) in the semi-discrete scheme (8), then one obtains a scheme that is indeed stable but such that the discrete operator involved in (9) becomes non block-diagonal. In particular, there is a priori no chance of deriving an explicit version of it. Second, it means that if one performs first the formal derivation (9)-(11) and only then discretizes (11) with (37) as is done in [28], then one obtains a scheme that is indeed explicit but that is not equivalent with the original formulation (18). In particular, the stability analysis developed in Section 2.3 does not apply to the resulting scheme. We finally note that the difference between the averaged vector product and the clustered one is proportional to the local gradient of the density, so that the difference between the two schemes becomes significant when turbulences are present.*

### 3.4. Detailed explicit formulas

We conclude this section by writing down the detailed formulas that result from the above procedure. To this end we assume that we have chosen an arbitrary triplet  $(\alpha, \beta, \gamma) \in \{-1, +1\}^3$ , and to fix the ideas we consider staggered grid operators  $\mathbf{b} \wedge_h$  and  $S(u)$  defined respectively by (38) with

$$\mathbf{b}^c|_{i,j,k} := \mathbf{b}|_{i,j,k} = \mathbf{b}(\mathbf{x}_{i,j,k})$$

and by (16) with

$$u^e|_{i,j+\frac{\beta}{2},k+\frac{\gamma}{2}} = u^e|_{i+\frac{\alpha}{2},j,k+\frac{\gamma}{2}} = u^e|_{i+\frac{\alpha}{2},j+\frac{\beta}{2},k} := u|_{i,j,k} = u(\mathbf{x}_{i,j,k}).$$

For conciseness we now denote  $\tilde{i} = i + \frac{\alpha}{2}$ ,  $\tilde{j} = j + \frac{\beta}{2}$  and  $\tilde{k} = k + \frac{\gamma}{2}$ . With the above choices the staggered grid operators correspond to

$$\begin{cases} (\mathbf{b} \wedge_h \mathbf{V}_h)_x|_{i,\tilde{j},\tilde{k}} := b_y|_{i,j,k} V_z|_{\tilde{i},\tilde{j},k} - b_z|_{i,j,k} V_y|_{\tilde{i},\tilde{j},k} \\ (\mathbf{b} \wedge_h \mathbf{V}_h)_y|_{i,\tilde{j},\tilde{k}} := b_z|_{i,j,k} V_x|_{\tilde{i},\tilde{j},k} - b_x|_{i,j,k} V_z|_{\tilde{i},\tilde{j},k} \\ (\mathbf{b} \wedge_h \mathbf{V}_h)_z|_{i,\tilde{j},\tilde{k}} := b_x|_{i,j,k} V_y|_{\tilde{i},\tilde{j},k} - b_y|_{i,j,k} V_x|_{\tilde{i},\tilde{j},k}. \end{cases} \quad (49)$$

and

$$\begin{cases} (S(u) \mathbf{V}_h)_x|_{i,\tilde{j},\tilde{k}} = u|_{i,j,k} V_x|_{i,\tilde{j},\tilde{k}} \\ (S(u) \mathbf{V}_h)_y|_{i,\tilde{j},\tilde{k}} = u|_{i,j,k} V_y|_{i,\tilde{j},\tilde{k}} \\ (S(u) \mathbf{V}_h)_z|_{i,\tilde{j},\tilde{k}} = u|_{i,j,k} V_z|_{i,\tilde{j},\tilde{k}} \end{cases} \quad (50)$$

and from Propositions 3.2 and 3.3 we know that the hypotheses of Proposition 3.4 are then satisfied. However we note that other choices could be made for  $\mathbf{b}^c$  and  $u^e$ . Thus, in the explicit

scheme (43) one first computes an auxiliary vector field  $\mathbf{W}_h^n$  given by

$$\begin{cases} W_x^n|_{i,j,\tilde{k}} = \Delta t \varepsilon_0 \omega_p^2 |_{i,j,k} E_x^n|_{i,j,\tilde{k}} + J_x^{n-\frac{1}{2}}|_{i,j,\tilde{k}} \\ \quad + \frac{\Delta t}{2} \omega_c |_{i,j,k} (b_y |_{i,j,k} J_z^{n-\frac{1}{2}}|_{i,j,\tilde{k}} - b_z |_{i,j,k} J_y^{n-\frac{1}{2}}|_{i,j,\tilde{k}}) \\ W_y^n|_{i,j,\tilde{k}} = \Delta t \varepsilon_0 \omega_p^2 |_{i,j,k} E_y^n|_{i,j,\tilde{k}} + J_y^{n-\frac{1}{2}}|_{i,j,\tilde{k}} \\ \quad + \frac{\Delta t}{2} \omega_c |_{i,j,k} (b_z |_{i,j,k} J_x^{n-\frac{1}{2}}|_{i,j,\tilde{k}} - b_x |_{i,j,k} J_z^{n-\frac{1}{2}}|_{i,j,\tilde{k}}) \\ W_z^n|_{i,j,\tilde{k}} = \Delta t \varepsilon_0 \omega_p^2 |_{i,j,k} E_z^n|_{i,j,\tilde{k}} + J_z^{n-\frac{1}{2}}|_{i,j,\tilde{k}} \\ \quad + \frac{\Delta t}{2} \omega_c |_{i,j,k} (b_x |_{i,j,k} J_y^{n-\frac{1}{2}}|_{i,j,\tilde{k}} - b_y |_{i,j,k} J_x^{n-\frac{1}{2}}|_{i,j,\tilde{k}}) \end{cases} \quad (51)$$

and expanding the formula  $\mathbf{J}_h^{n+\frac{1}{2}} = \left( \mathbf{I} + S\left(\frac{\theta}{1+\theta^2}\right)(\mathbf{b} \wedge_h) + S\left(\frac{\theta^2}{1+\theta^2}\right)(\mathbf{b} \wedge_h)^2 \right) \mathbf{W}_h^n$  for the current, one obtains

$$\begin{cases} J_x^{n+\frac{1}{2}}|_{i,j,\tilde{k}} = \left(1 - \left(\frac{\theta^2}{1+\theta^2}\right)\right) |_{i,j,k} (b_y^2 - b_z^2) |_{i,j,k} W_x^n|_{i,j,\tilde{k}} \\ \quad + \left(\left(\frac{\theta}{1+\theta^2}\right) |_{i,j,k} (-b_z + \theta b_x b_y) |_{i,j,k}\right) W_y^n|_{i,j,\tilde{k}} \\ \quad + \left(\left(\frac{\theta}{1+\theta^2}\right) |_{i,j,k} (b_y + \theta b_x b_z) |_{i,j,k}\right) W_z^n|_{i,j,\tilde{k}} \\ J_y^{n+\frac{1}{2}}|_{i,j,\tilde{k}} = \left(1 - \left(\frac{\theta^2}{1+\theta^2}\right)\right) |_{i,j,k} (b_z^2 - b_x^2) |_{i,j,k} W_y^n|_{i,j,\tilde{k}} \\ \quad + \left(\left(\frac{\theta}{1+\theta^2}\right) |_{i,j,k} (-b_x + \theta b_y b_z) |_{i,j,k}\right) W_z^n|_{i,j,\tilde{k}} \\ \quad + \left(\left(\frac{\theta}{1+\theta^2}\right) |_{i,j,k} (b_z + \theta b_y b_x) |_{i,j,k}\right) W_x^n|_{i,j,\tilde{k}} \\ J_z^{n+\frac{1}{2}}|_{i,j,\tilde{k}} = \left(1 - \left(\frac{\theta^2}{1+\theta^2}\right)\right) |_{i,j,k} (b_x^2 - b_y^2) |_{i,j,k} W_z^n|_{i,j,\tilde{k}} \\ \quad + \left(\left(\frac{\theta}{1+\theta^2}\right) |_{i,j,k} (-b_y + \theta b_z b_x) |_{i,j,k}\right) W_x^n|_{i,j,\tilde{k}} \\ \quad + \left(\left(\frac{\theta}{1+\theta^2}\right) |_{i,j,k} (b_x + \theta b_z b_y) |_{i,j,k}\right) W_y^n|_{i,j,\tilde{k}}. \end{cases} \quad (52)$$

In the scheme (43), one computes an auxiliary vector field  $\mathbf{W}_h^n$  given by a formula similar to (51), specifically

$$\begin{cases} W_x^n|_{i,j,\tilde{k}} = \frac{1}{\eta|_{i,j,k}} \left( \Delta t \varepsilon_0 \omega_p^2 |_{i,j,k} (E_x^n|_{i,j,\tilde{k}} + \frac{\Delta t}{2\varepsilon_0} (R\mathbf{H}_h^{n+\frac{1}{2}})_x |_{i,j,\tilde{k}}) \right. \\ \quad + \left(1 - \frac{\Delta t^2}{4} \omega_p^2 |_{i,j,k}\right) J_x^n|_{i,j,\tilde{k}} \\ \quad + \frac{\Delta t}{2} \omega_c |_{i,j,k} (b_y |_{i,j,k} J_z^n|_{i,j,\tilde{k}} - b_z |_{i,j,k} J_y^n|_{i,j,\tilde{k}}) \Big) \\ W_y^n|_{i,j,\tilde{k}} = \frac{1}{\eta|_{i,j,k}} \left( \Delta t \varepsilon_0 \omega_p^2 |_{i,j,k} (E_y^n|_{i,j,\tilde{k}} + \frac{\Delta t}{2\varepsilon_0} (R\mathbf{H}_h^{n+\frac{1}{2}})_y |_{i,j,\tilde{k}}) \right. \\ \quad + \left(1 - \frac{\Delta t^2}{4} \omega_p^2 |_{i,j,k}\right) J_y^n|_{i,j,\tilde{k}} \\ \quad + \frac{\Delta t}{2} \omega_c |_{i,j,k} (b_z |_{i,j,k} J_x^n|_{i,j,\tilde{k}} - b_x |_{i,j,k} J_z^n|_{i,j,\tilde{k}}) \Big) \\ W_z^n|_{i,j,\tilde{k}} = \frac{1}{\eta|_{i,j,k}} \left( \Delta t \varepsilon_0 \omega_p^2 |_{i,j,k} (E_z^n|_{i,j,\tilde{k}} + \frac{\Delta t}{2\varepsilon_0} (R\mathbf{H}_h^{n+\frac{1}{2}})_z |_{i,j,\tilde{k}}) \right. \\ \quad + \left(1 - \frac{\Delta t^2}{4} \omega_p^2 |_{i,j,k}\right) J_z^n|_{i,j,\tilde{k}} \\ \quad + \frac{\Delta t}{2} \omega_c |_{i,j,k} (b_x |_{i,j,k} J_y^n|_{i,j,\tilde{k}} - b_y |_{i,j,k} J_x^n|_{i,j,\tilde{k}}) \Big). \end{cases} \quad (53)$$

The current field  $\mathbf{J}_h^{n+1}$  is then given by the same formula as before, i.e., (52) applies again to compute the entries  $J_x^{n+1}|_{i,\tilde{j},\tilde{k}}$ ,  $J_y^{n+1}|_{i,\tilde{j},\tilde{k}}$  and  $J_z^{n+1}|_{i,\tilde{j},\tilde{k}}$ .

For the sake of completeness we also recall the classical formulas corresponding to the matrix (15) that approximates the curl operator between the spaces  $\mathcal{V}_f$  and  $\mathcal{V}_e$  (the adjoint curl operator being approximated by the transpose matrix).

$$\left\{ \begin{array}{l} (\mathbf{R}\mathbf{H}_h^{n+\frac{1}{2}})_x|_{i,j+\frac{1}{2},k+\frac{1}{2}} = \frac{1}{h}(H_z^{n+\frac{1}{2}}|_{i,j+1,k+\frac{1}{2}} - H_z^{n+\frac{1}{2}}|_{i,j,k+\frac{1}{2}}) \\ \quad - \frac{1}{h}(H_y^{n+\frac{1}{2}}|_{i,j+\frac{1}{2},k+1} - H_y^{n+\frac{1}{2}}|_{i,j+\frac{1}{2},k}) \\ (\mathbf{R}\mathbf{H}_h^{n+\frac{1}{2}})_y|_{i+\frac{1}{2},j,k+\frac{1}{2}} = \frac{1}{h}(H_x^{n+\frac{1}{2}}|_{i+\frac{1}{2},j,k+1} - H_x^{n+\frac{1}{2}}|_{i+\frac{1}{2},j,k}) \\ \quad - \frac{1}{h}(H_z^{n+\frac{1}{2}}|_{i+1,j,k+\frac{1}{2}} - H_z^{n+\frac{1}{2}}|_{i,j,k+\frac{1}{2}}) \\ (\mathbf{R}\mathbf{H}_h^{n+\frac{1}{2}})_z|_{i+\frac{1}{2},j+\frac{1}{2},k} = \frac{1}{h}(H_y^{n+\frac{1}{2}}|_{i+1,j+\frac{1}{2},k} - H_y^{n+\frac{1}{2}}|_{i,j+\frac{1}{2},k}) \\ \quad - \frac{1}{h}(H_x^{n+\frac{1}{2}}|_{i+\frac{1}{2},j+1,k} - H_x^{n+\frac{1}{2}}|_{i+\frac{1}{2},j,k}). \end{array} \right. \quad (54)$$

#### 4. Numerical results

Our numerical tests are relevant for reflectometry studies in fusion plasmas or ionospheric probing. The domain of computation is 2D and we consider X-mode configuration, meaning that polarization is transverse electric as in (6). The 2D implementation maintains all the difficulties studied in this work and is simpler to implement and validate. We present three test problems on time-independent plasmas which are the main subject of this work, together with one additional time-dependent plasma case, a subject treated in Appendix A. The tests have been conducted within the REFMULX code, a 2D FDTD X-mode code developed at the IPFN (Instituto de Plasmas e Fusão Nuclear) [8]. Simulations were done on a grid of  $1500 \times 1000$  points with launched frequency  $f = 35$  GHz (that is in terms of vacuum wavelengths  $65.6\lambda_0 \times 43.8\lambda_0$ , given a spatial resolution of  $h = \Delta x = \Delta y = \lambda_0/20$ , where  $\lambda_0$  is the the vacuum wavelength at 40 GHz). Figure 3 shows that the coefficients of the problem (in terms of  $f_c = \frac{\omega_c}{2\pi}$  and  $f_p = \frac{\omega_p}{2\pi}$ ) are of the same order.

##### 4.1. Implementation details

The former implementation of REFMULX, version 1.0 (v1.0) [8], used a kernel referred to as the Xu-Yuan Kernel (XYK) whose original description appears in [28]. We use it in some of the tests to show its limitations in the context of this work.

The new version of REFMULX, version 2.0 (v2.0), includes two kernels: (i) the explicit scheme (42) equivalent to (18), which corresponds to a rewriting of the XYK scheme where the symmetric vector products (37) have been substituted by the non-centered formulas (38)-(39) (with a given arbitrary translation pattern fixed once and for all) ; and (ii) the new conservative scheme (43) equivalent to (32), proposed in this work to improve the CFL constraint.

The CFL condition has been checked to satisfy the numerical formula (28) which is enough to get the numerical stability in the energy norm, but can be also considered as a physical evaluation of the speed of numerical waves.

#### 4.2. Tests

The tests use physically based coefficients.

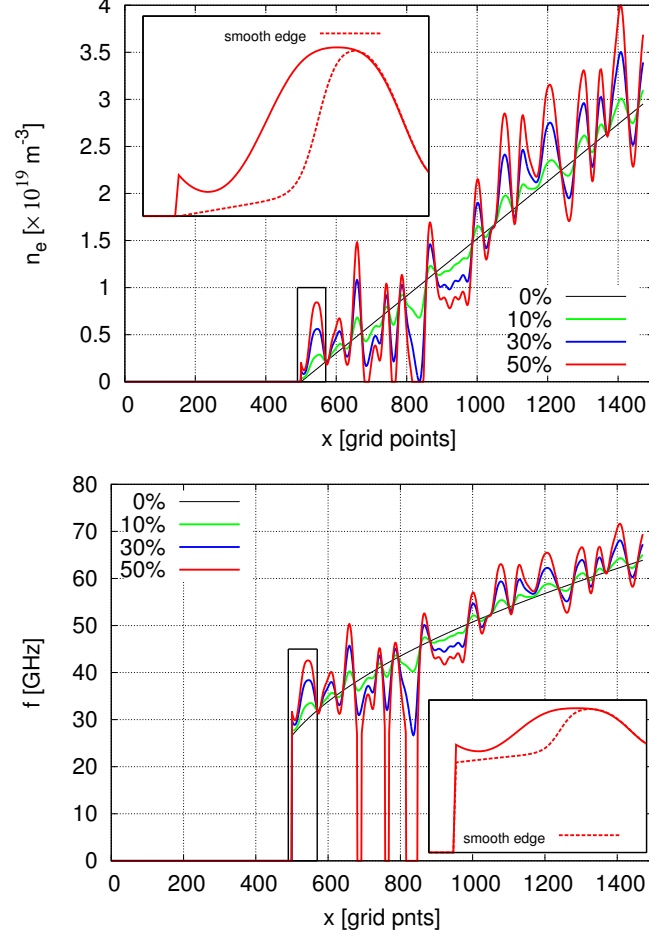


Figure 3: On top, the plasma background density  $N_{e0}(x)$  (0%) used for the non turbulent tests is shown together with turbulent densities  $N_e(x, y) = N_{e0}(x) + \delta N_e(x, y)$  for three levels of turbulence  $\delta N_e(x_{35})/N_e(x_{35})$  of 10%, 30% and 50% RMS (here  $x_{35}$  is the position of the cut-off for a probing wave launched at 35 GHz). On bottom, the corresponding upper cutoff frequencies  $f_U = (4\pi)^{-1}((\omega_p(x)^2 + \omega_c^2)^{1/2} + \omega_c)$  (above the cyclotron frequency  $f_c = \frac{\omega_c}{2\pi} \approx 26.6 \text{ GHz}$ ) are presented for a magnetic field of  $B_0 = 0.95 \text{ T}$ . In zones where  $n_e = 0$  (meaning no plasma), we arbitrarily set  $f_U = 0$  thus following the physical intuition. This is visible in the bottom figure where  $f_U$  vanishes in the same zones than  $n_e$ . The insets show the sharp discontinuities at the left plasma edge, and the smooth transitions used in some of the test cases.

The plasma setup has a background density denoted as  $N_{e0}(x)$  which increases linearly along  $x$  (radial direction), with its edge at  $x = 500$  grid points ( $21.9\lambda_{35 \text{ GHz}}$ ). Turbulence  $\delta N_e(x, y)$  is added to the background plasma  $N_e(x, y) = N_{e0}(x) + \delta N_e(x, y)$ . In Figure 3 a typical plasma density profile is depicted in the top figure where the electronic density is plotted against the horizontal axis (measured in cell points knowing that  $h = \Delta x = \Delta y = \lambda_0/20$ , where  $\lambda_0$  is the

vacuum wavelength at 40 GHz). The background density profile  $N_{e_0}(x)$  (without turbulence, i.e. 0%) is used in the simulations described in 4.2.1–4.2.2. Simulation tests on 4.2.3–4.2.4 are done with turbulent plasma with levels  $\delta N_e(x_{35})/N_e(x_{35})$  ranging from 5%–50% in RMS: here  $x_{35}$  is the position of the cut-off for a probing wave launched at 35 GHz. In Figure 3 density profiles at different turbulence levels 10%, 30% and 50% are shown. In the inset a zoom of the edge density (only for the 50% curve, for clarity) is shown. Its meaning will be explained in 4.2.3. The bottom of Figure 3 shows the upper cut-off frequency  $f_U$  for a constant external magnetic field of  $B_0 = 0.95T$ : the upper cut-off frequency is above the electronic cyclotron frequency  $f_c = \frac{\omega_c}{2\pi} \approx 26.6$  GHz. Be careful that for physical reasons, we decided arbitrarily that a vanishing electronic density  $n_e$  corresponds to  $f_U = 0$ .

The first test presents a comparison of (42) and (43) to verify that the pseudo-energy is conserved.

In the second set of tests very long simulations are performed to show the influence of round-off errors on the numerical stability/instability when second order symmetric formulas like (37) are used. Here there are no density fluctuations, so the problem is without turbulence.

After these first numerical tests, the following ones are more closely connected to scenarios of interest for the fusion plasma community. Different types of turbulence have been tested to determine if the numerical stability is affected by the turbulence properties. As Gaussian wave-number spectrum is representative of the turbulence behavior, this kind of turbulence is used in all the following tests. The different test cases appear from the simplest to the more general one.

In the third family of tests we consider steady state turbulent cases and we assess the improved stability offered by the clustered (non-centered) vector-product (38). Results are shown for different levels of turbulence, and we specify some of the parameters used to stabilize the computations over simulation times representing at least one million time steps.

The fourth tests address an extension of the proposed scheme in order to assess the stability for time-dependent turbulence-wave interaction simulations, which are closer to experiments. There the turbulence has a temporal dependency  $N_e(x, y, t) = N_{e_0}(x) + \delta N_e(x, y, t)$ , being able to evolve with a velocity along  $y$  (poloidal) and  $x$  (radial) directions.

#### 4.2.1. Test 1: pseudo-energy conservation

Here we compare the results obtained with the new explicit schemes (42) and (43), both implemented with the clustered (non-centered) vector product (38). This test intends to verify the theoretical conservation properties in the numerical results for an extremely steep initial signal. The initial data is a Dirac mass, and we represent in figure 4 the 2D version of the pseudo-energy (31) versus the time steps. The Dirac mass is located 100 grid cells inside the plasma, therefore being at  $i_s = 600$  where  $i_s = 500$  is the beginning of the electronic density ramp, see Figure 3 (curve 0%). The external magnetic field used in this set of runs was  $B_0 = 0.95T$ .

At the position  $(i_s, j_s) = (600, 500)$ , an impulse is excited at iteration  $n = 0$ , with a value  $H_z(i_s, j_s, 0) = \frac{1}{\mu_0} \frac{1}{\sqrt{\Delta x \Delta y}}$  to have a energy density of  $\frac{1}{\Delta x \Delta y}$  resulting in an initial pseudo-energy  $\mathcal{E}_h^0 = 1$ , see (31). The number of iterations considered is  $n = 700$ , not long enough to allow a response from the boundary zone. Results appear in Figure 4, where the exact preservation of the pseudo-energy by the scheme (43) confirms the theory, see Proposition 2.3. We also note that according to Proposition 2.1, the pseudo-energy should be exactly preserved by the scheme (42).

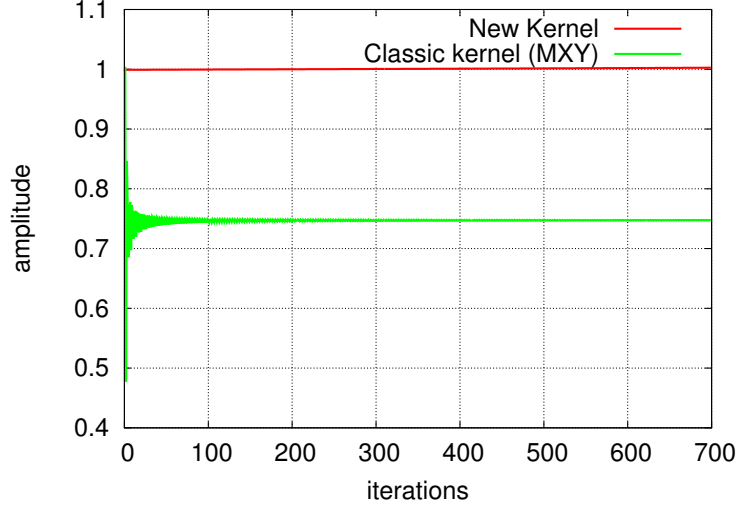


Figure 4: Pseudo-energy (31) of the numerical solutions computed with the schemes (42) (in green) and (43) (in red), starting from an initial Dirac impulse far from the boundaries. The exact preservation of the pseudo-energy by (43) is in accordance with the theory, see Proposition 2.3.

#### 4.2.2. Test 2: stability of wave propagation (without turbulence)

In this series of tests we assess the long time stability of the code in the absence of turbulence. Thus the same linear plasma density  $N_{e0}$  is considered as in the previous test cases (cf. Figure 3, curve 0%), and a wave is launched with frequency  $f = 35$  GHz from the left part of the domain, as in Figure 8. Here the standard XYK scheme (involving averaged vector products and implemented in REFMULX v1.0) shows to be stable up to 100 thousand iterations but for longer runs is unstable. By contrast, on REFMULX v2.0 both kernels (42) and (43) show an augmented stability up to 3.5 million iterations. Increasing further the duration of the runs we have verified that they became unstable when using FLOAT. Written in DOUBLE the new kernels reached 8 million iterations in our trials. The natural interpretation is that the control of roundoff errors is important for such difficult problems. In Figure 5 we present the longest run tried in these tests, using a logarithmic scale to better visualize the exponential growth of the unstable run.

#### 4.2.3. Test 3: wave propagation in steady state turbulent plasma

The aim of this part is to detail some points on the stability domain of the computation of X-mode propagation in turbulent plasma. We start with the standard scheme XYK [28], used on v1.0 (with averaged vector products), and further ahead use both the schemes (42) and (43) implemented in v2.0.

The original XYK, which was already seen to be unstable without turbulence for simulation times of about 100 thousand iterations, shows very low stability properties when adding high levels of turbulence. In fact, even before 100 thousand iterations the runs are seldom stable. The characterization of these instabilities with turbulence has some interest to better understand the phenomenon. In Figure 6 we present the results obtained with XYK (REFMULX v1.0), using a logarithmic scale to better visualize the exponential growth of the instabilities. In the top part

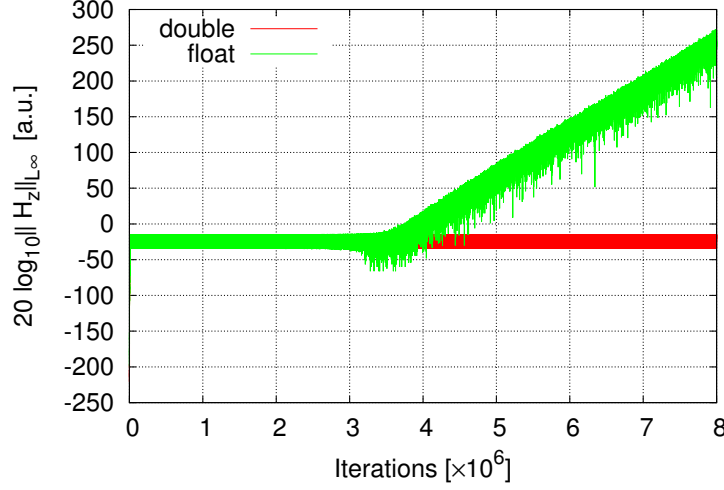


Figure 5: Amplitude of the solution  $H_z$  computed with the new scheme (42) implemented in REFMULX v2.0 versus iteration number for the non-turbulent test case described in Section 4.2.2. Here the use of FLOAT results in a run that is unstable after  $\approx 3.5$  million time steps, and a logarithmic scale is used to better visualize the exponential growth. Using DOUBLE the run remains stable during the 8 million time steps of the complete simulation. Similar results were obtained with the new scheme (43) and are not shown here.

of Figure 6 we show the runs obtained using turbulent plasma density profiles where on the left plasma edge a sharp discontinuity is present, as seen in Figure 3 for three levels of turbulence. In the lower part of Figure 6 we show the runs obtained by replacing those sharp discontinuities by a smooth transition, as seen in the inset of Figure 3 (only for the 50% case, for clarity). Although very localized, this slight modification of  $N_e$  has a considerable impact on the overall results: for instance, with a smooth plasma edge the 10% case becomes *stable*.

In Figure 7 we next represent the time steps where the instabilities are triggered (upper part) together with their growth rates (lower part) for several levels of turbulence. We thus verify that in the presence of a sharp discontinuity at the plasma edge, the instabilities are triggered sooner and they grow faster than in the case where the transition is smooth. And in both cases we observe that the same occurs (earlier instabilities and higher growth rates) as the turbulence amplitude increases. Note that these results are to be looked at as a behavior guideline, not as an absolute value: indeed, using a different turbulence matrix profile  $\delta N_e$  with same spectrum we may very well obtain an unstable run in the 10% case with a smooth plasma edge.

In Figures 8 and 9 we next show numerical results obtained with the new explicit scheme (43) involving non-centered vector products, as implemented in REFMULX v2.0. Here we essentially observe that all the problems that were unstable before become stable, and that the spurious exponential growths illustrated in Figure 6 are prevented. This confirms in practice the theoretical stability result proved in Proposition 3.4, in particular this shows that in this case the presence of boundary conditions does not spoil the numerical stability of the scheme. In Figure 8 we represent a typical snapshot that is representative of all these stable runs. There the field  $H_z$  is plotted at iteration 80,000 for a turbulence level of 10%, together with the distorted contours

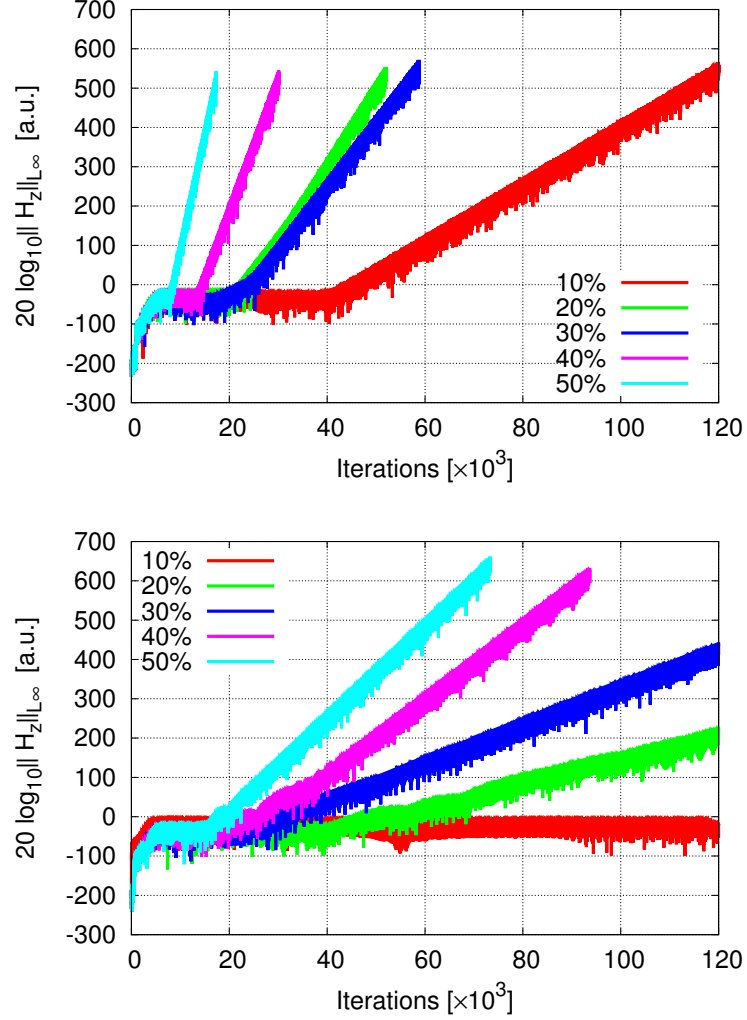


Figure 6: Return signals  $20 \log_{10} \|H_z\|_{L^\infty}$  for wave simulations in a steady state turbulent plasma using the XYK scheme implemented in REF-MULX v1.0 with a sharp edge (top) and a smooth transition (bottom) in the plasma density profile, see Figure 3 and details in the text.

of the upper cutoff  $f_U = 35$  GHz (defined in Figure 3). In Figure 9 we next show the runs obtained with the scheme (43) with different levels of turbulence and a sharp plasma edge (see Figure 3). When compared with the runs obtained with the centered vector product, the result is outstanding in terms of stability. Note that simulations are now stable at least up to 1 million iterations. Results with the other new scheme (42) present the same level of stability, and are not shown here.

As we have identified that the spatio-temporal evolution (stability) is highly connected to the



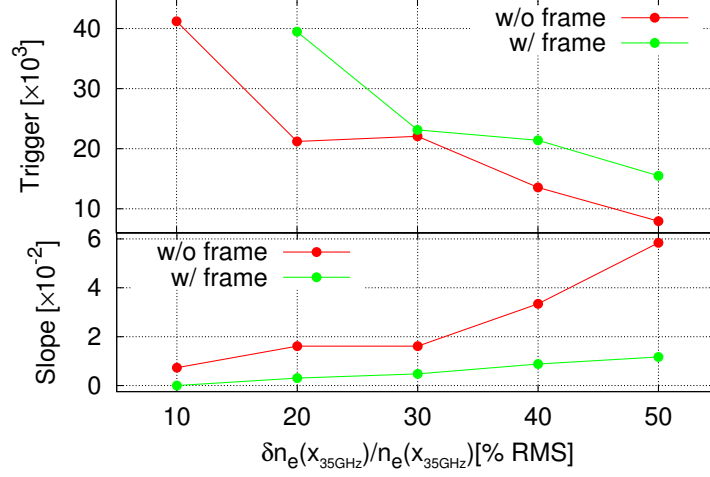


Figure 7: On top the trigger instants of the instabilities presented in Figure 6 and on bottom their growth rates. Here the framed cases correspond to a smooth transition at the plasma edge, see Figure 3.

behavior to sharp density gradients at the plasma edge, the use of a smoothed density fluctuations profile can play a role on stability improvement. Another important point when implementing the kernels is to optimize the numerical operations by minimizing the number of operations, especially divisions, which not only increases performance but also improves stability for such challenging problems. This improvement is enhanced when one switches from single to double precision, extending the validity domain of both the new kernels (42) and (43) to more than one million iterations. This number is enough to treat cases relevant for fusion reactors or ITER. However if one wants to study the probing beam widening induced by density fluctuations [24], the requirements are even higher. The turbulence profile  $\delta N_e$  needs to move so that the wave intensity can be averaged over a number of time steps, and we are closer to the limits of HPC machines. This motivates the following section, devoted to time-dependent cases.

#### 4.2.4. Test 4: time-dependent plasma

Although not the main subject of this article, the case of a time-dependent plasma density is addressed here because of its importance on some problems in plasma physics. Thus, our objective is not to provide a full treatment of this issue, which we divert to a later study, but to alert to some of the problems it poses, namely as stability is concerned. In the time-dependent case indeed an additional term  $\partial_t N_e / N_e$  is introduced, and our numerical scheme needs to be modified as described in Appendix A. As a consequence, we observe that numerical problems can arise at the plasma edge where the plasma density  $N_e$  can be very low and the variations  $\partial_t N_e$  very high. Again, this can induce numerical instabilities. Moreover this problem is more difficult to analyze than before since the total energy is no more conserved, due to the source term proportional to  $\partial_t N_e$ , see (A.5) below.

To reduce this potential instability effect, we will use some heuristics deduced from the previous tests. Indeed one can use the *smoothed* edge plasma (see Figure 3) to reduce the local density

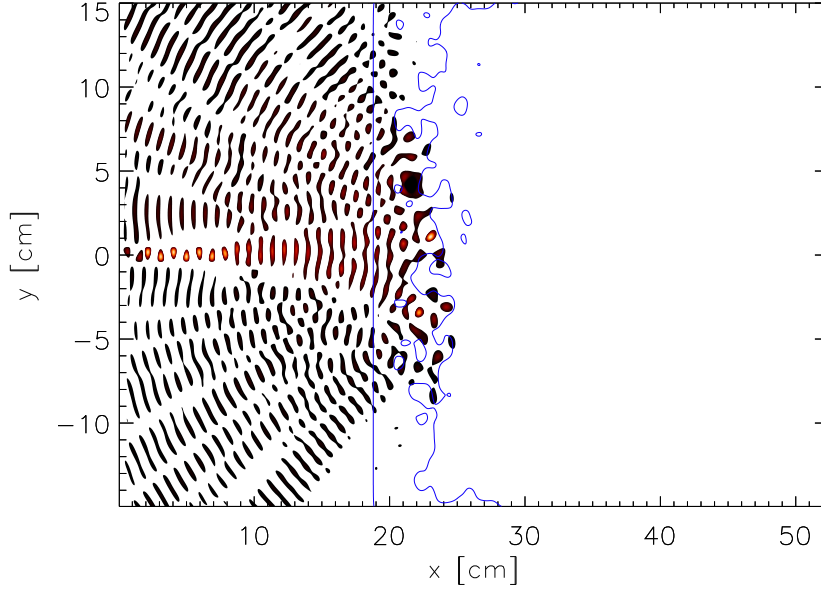


Figure 8: Snapshot of the positive part of the magnetic field, that is  $\max(H_z, 0)$ , taken at iteration 80,000 with a turbulence level of 10%, using the new scheme (43). The distorted contour lines corresponds to the upper cutoff  $f_U = 35$  GHz as defined in the caption of Figure 3.

gradient and to regularize the plasma density profile. We introduce two directions in which the turbulence matrix  $\delta N_e$  moves. One is parallel to the plasma edge and it is called poloidal (using tokamak terminology as a reference) and the other is co-linear to the plasma density gradient and is the radial direction. Different studies have been done as a function of the turbulence level, the direction, and velocity of the motion. The movement of the density is done sliding the profile matrix  $\delta N_e$  along the required direction (poloidally or radially), shifting the lines or columns within the discretization grid. The velocity of a fusion plasma is much slower than the time frame of the electromagnetic wave ( $V \ll c$ ), with top values in the order of km/s. To speed up simulations we allow the plasma to move faster than the real velocities, correcting the results to accommodate for the higher simulation speed [6]. Nevertheless, there are limits to that scaling and the plasma velocity must be kept at a fraction of the speed of light to maintain physical meaning [5],[6], usually around 2% of  $c$ , always less than 10%. Therefore the number of iterations for a  $\delta_{x,y}$  displacement of the matrix  $\delta N_e$  to occur is  $\Delta n = 2(V/c)^{-1}$ . While ensuring the correct velocity, this shift of the matrix every  $\Delta n$  iterations involves a sudden variation in the value of  $N_e$  at each point. A better method is to check at each iteration what fraction of  $\Delta n$  has already elapsed and set  $N_e$  as a *weight* of the present density and a  $(1 - \text{weight})$  of the next value, securing a much smoother motion with the same velocity and showing much cleaner spectra. The movement of the turbulence matrix  $\delta N_e$  reduces the stability and the correction  $\partial_t N_e / N_e$  limits it even further. In all tests the time dependence reduces severely the level of turbulence were the code holds its stability as well as the duration of the stable runs. Introducing a poloidal movement with  $V/c = 1.4\%$  on

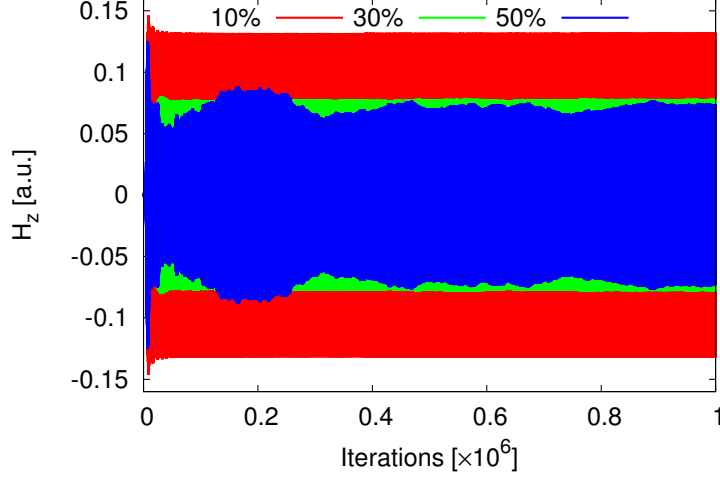


Figure 9: Return signals are shown in linear scale for different levels of turbulence using the new scheme (43) implemented in REFMULX v2.0 which involves the clustered (non-centered) vector product (38). Here every case is stable, although with more amplitude for a high level of turbulence. Note that the number of time steps is much greater than in Figure 6 to illustrate the improved stability of the method. Runs obtained with the scheme (42) also implemented in REFMULX v2.0 show the same level of stability and are not presented here.

the simulation set-up used in the previous tests, neglecting the correction  $\partial_t N_e / N_e$ , a stable run of 140 thousand iterations is obtained for a turbulence level of 10%. Including the correction  $\partial_t N_e / N_e$ , the maximum turbulence level drops to 5%. Preliminary tests with radial displacement of the plasma seem to be more stable than the poloidal case when neglecting  $\partial_t N_e / N_e$  but the degradation in stability brought by the correction is much higher than in the poloidal case.

Partial conclusions to have a stable code in the time-dependent case are: reduce the turbulence level at the edge, be sure that turbulent matrix  $\delta N_e$  has no discontinuity inside, optimize the numerical operation avoiding division, use double precision.

## 5. Conclusion

We have detailed the influence of strong gradients in the plasma parameters on the stability in the energy norm of a Yee scheme coupled with a linear current model. The main asset is the identification of an instability phenomenon triggered by strong gradients of  $N_e$ . This is related to the use of a centered evaluation of the vector product. Different solutions have been proposed to get a scheme guaranteed stable in energy norm, all of them being related to the use of a clustered vector product acting on edgewise vector fields that satisfies two compatibility relations. The first one,

$$\langle \mathbf{V}_h, \mathbf{b} \wedge_h \mathbf{V}_h \rangle_h = 0 \quad \text{for } \mathbf{V}_h \in \mathcal{V}_e,$$

is in our context necessary to get a scheme stable in the energy norm. The second one reads

$$(\mathbf{b} \wedge_h)^3 \mathbf{V}_h = -(\mathbf{b} \wedge_h) \mathbf{V}_h \quad \text{for } \mathbf{V}_h \in \mathcal{V}_e$$

and expresses the fact that, as an operator acting on edgewise vector fields,  $\mathbf{b} \wedge_h$  is *block-diagonal*. As such, it allows to get an explicit scheme with the same CFL condition as in vacuum. Both conditions are very natural since they are true of course for the exact operator, given that  $|\mathbf{b}| = 1$ . Our numerical tests confirm the theoretical analysis and they assess the remarkable stability property of the compact first order vector product. Further research will be devoted to the study of moving plasmas for which  $\partial_t N_e \neq 0$  and to full 3D implementations involving all wave components. These are of main interest to treat cases closer to experiments where the full description should include turbulence and plasma motions. This stage is a path to the introduction of synthetic diagnostics in turbulence codes to predict signals coming from either reflectometer signals or scattering diagnostics. Overall, we emphasize that the applications of the numerical schemes introduced in this article are wider than those presented. They include the beam widening of electron cyclotron heating as well as other heating system using waves, space physics or telecommunications.

### Appendix A. Time dependent plasma densities

If the electronic density  $N_e$  varies with time, then the current density satisfies

$$\partial_t \mathbf{J} = e N_e \partial_t \mathbf{u}_e + \mathbf{J} \frac{\partial_t N_e}{N_e}$$

and using  $\frac{\partial_t N_e}{N_e} = 2 \frac{\partial_t \omega_p}{\omega_p}$  we see that (7) becomes

$$\begin{cases} -\varepsilon_0 \partial_t \mathbf{E} + \text{curl } \mathbf{H} = \mathbf{J}, \\ \mu_0 \partial_t \mathbf{H} + \text{curl } \mathbf{E} = 0, \\ \partial_t \mathbf{J} = \varepsilon_0 \omega_p^2 \mathbf{E} + \omega_c \mathbf{b} \wedge \mathbf{J} + 2 \mathbf{J} \frac{\partial_t \omega_p}{\omega_p}. \end{cases} \quad (\text{A.1})$$

Turning to the symmetrized variables  $\widehat{\mathbf{E}} = \frac{\mathbf{E}}{c}$ ,  $\widehat{\mathbf{J}} = \frac{\mathbf{J}}{\omega_p c \varepsilon_0}$  and  $\widehat{\mathbf{B}} = \mathbf{B} = \mu_0 \mathbf{H}$  we observe that the time-varying plasma density yields  $\partial_t \widehat{\mathbf{J}} = \frac{\partial_t \mathbf{J}}{\omega_p c \varepsilon_0} - \frac{\partial_t \omega_p}{\omega_p} \widehat{\mathbf{J}}$ . We thus obtain

$$\begin{cases} \partial_t \widehat{\mathbf{E}} - c \text{curl } \widehat{\mathbf{B}} = -\omega_p \widehat{\mathbf{J}} \\ \partial_t \widehat{\mathbf{B}} + c \text{curl } \widehat{\mathbf{E}} = 0 \\ \partial_t \widehat{\mathbf{J}} = \omega_p \widehat{\mathbf{E}} + \omega_c \mathbf{b} \wedge \widehat{\mathbf{J}} + \frac{\partial_t \omega_p}{\omega_p} \widehat{\mathbf{J}}. \end{cases} \quad (\text{A.2})$$

Discretizing this system as in (30) yields then

$$\begin{cases} \frac{\widehat{\mathbf{E}}_h^{n+1} - \widehat{\mathbf{E}}_h^n}{\Delta t} = c R \widehat{\mathbf{B}}_h^{n+\frac{1}{2}} - S(\omega_p^{n+\frac{1}{2}}) \widehat{\mathbf{J}}_h^{n+\frac{1}{2}} \\ \frac{\widehat{\mathbf{B}}_h^{n+\frac{1}{2}} - \widehat{\mathbf{B}}_h^{n-\frac{1}{2}}}{\Delta t} = -c R^t \widehat{\mathbf{E}}_h^n \\ \frac{\widehat{\mathbf{J}}_h^{n+1} - \widehat{\mathbf{J}}_h^n}{\Delta t} = S(\omega_p^{n+\frac{1}{2}}) \widehat{\mathbf{E}}_h^{n+\frac{1}{2}} + \left( S(\omega_c)(\mathbf{b} \wedge_h) + S\left(\frac{\omega_p^{n+1} - \omega_p^n}{\Delta t \omega_p^{n+\frac{1}{2}}}\right) \right) \widehat{\mathbf{J}}_h^{n+\frac{1}{2}} \end{cases} \quad (\text{A.3})$$

where  $\widehat{\mathbf{E}}_h^{n+\frac{1}{2}} = \frac{1}{2}(\widehat{\mathbf{E}}_h^n + \widehat{\mathbf{E}}_h^{n+1})$  and  $\widehat{\mathbf{J}}_h^{n+\frac{1}{2}} = \frac{1}{2}(\widehat{\mathbf{J}}_h^n + \widehat{\mathbf{J}}_h^{n+1})$ . For this scheme the pseudo-energy

$$\mathcal{E}_h^n = \|\widehat{\mathbf{E}}_h^n\|_h^2 + \|\widehat{\mathbf{B}}_h^{n-\frac{1}{2}}\|_h^2 + \|\widehat{\mathbf{J}}_h^n\|_h^2 - \Delta t \langle \widehat{\mathbf{E}}_h^n, cR\widehat{\mathbf{B}}_h^{n-\frac{1}{2}} \rangle_h \quad (\text{A.4})$$

satisfies

$$\frac{\mathcal{E}_h^{n+1} - \mathcal{E}_h^n}{\Delta t} = \left\langle S\left(\frac{\omega_p^{n+1} - \omega_p^n}{\Delta t \omega_p^{n+\frac{1}{2}}}\right) \widehat{\mathbf{J}}_h^{n+\frac{1}{2}}, \widehat{\mathbf{J}}_h^{n+\frac{1}{2}} \right\rangle$$

which is a centered approximation of the variation of the exact energy (21),

$$\frac{d}{dt} \mathcal{E}(t) = \int_{\Omega} \frac{\partial_t \omega_p(t, \mathbf{x})}{2\mu_0 \omega_p(t, \mathbf{x})} |\widehat{\mathbf{J}}(t, \mathbf{x})|^2 d\mathbf{x}. \quad (\text{A.5})$$

Note that it is possible to rewrite the scheme (A.3) in the original variables. Because of the time dependence of  $\omega_p$ , the latter are given by  $\widehat{\mathbf{E}}_h^n = \frac{\mathbf{E}_h^n}{c}$ ,  $\widehat{\mathbf{J}}_h^n = S(\omega_p^n) \frac{\mathbf{J}_h^n}{c\varepsilon_0}$  and  $\widehat{\mathbf{B}}_h^{n+\frac{1}{2}} = \mu_0 \mathbf{H}_h^{n+\frac{1}{2}}$  which allows to rewrite (A.3) as

$$\left\{ \begin{array}{l} \varepsilon_0 \frac{\mathbf{E}_h^{n+1} - \mathbf{E}_h^n}{\Delta t} = R\mathbf{H}_h^{n+\frac{1}{2}} - \frac{1}{2} \left( S\left(\frac{\omega_p^{n+\frac{1}{2}}}{\omega_p^n}\right) \mathbf{J}_h^n + S\left(\frac{\omega_p^{n+\frac{1}{2}}}{\omega_p^{n+1}}\right) \mathbf{J}_h^{n+1} \right) \\ \mu_0 \frac{\mathbf{H}_h^{n+\frac{3}{2}} - \mathbf{H}_h^{n+\frac{1}{2}}}{\Delta t} = -R^t \mathbf{E}_h^{n+1} \\ \frac{\mathbf{J}_h^{n+1} - \mathbf{J}_h^n}{\Delta t} = \varepsilon_0 S(\omega_p^{n+\frac{1}{2}} \omega_p^{n+1}) \mathbf{E}_h^{n+\frac{1}{2}} + S\left(\frac{\omega_p^{n+1} - \omega_p^n}{\Delta t \omega_p^n}\right) \mathbf{J}_h^n \\ \quad + \left( S(\omega_c)(\mathbf{b} \wedge_h) + S\left(\frac{\omega_p^{n+1} - \omega_p^n}{\Delta t \omega_p^{n+\frac{1}{2}}}\right) \right) \widehat{\mathbf{J}}_h^{n+\frac{1}{2}} \end{array} \right. \quad (\text{A.6})$$

where  $\mathbf{E}_h^{n+\frac{1}{2}} = \frac{1}{2}(\mathbf{E}_h^n + \mathbf{E}_h^{n+1})$  and  $\mathbf{J}_h^{n+\frac{1}{2}} = \frac{1}{2} \left( S\left(\frac{\omega_p^{n+1}}{\omega_p^n}\right) \mathbf{J}_h^n + \mathbf{J}_h^{n+1} \right)$ . Clearly, this is not an intuitive discretization of (A.1).

## Appendix B. Resolution with time averaged operators

We here detail how the formulas in Section 3.3 are modified in the case where some averaging in time is introduced through the use of a time-dependent cluster pattern

$$(\alpha^n, \beta^n, \gamma^n) \in \{-1, +1\}^3, \quad n \in \mathbb{N}. \quad (\text{B.1})$$

In such a case we observe that in order to satisfy Properties (25), (33) and (40) the staggered grid operators should also change in time. Let us then denote by  $\mathbf{b} \wedge_h^n$  and  $S^n(u)$  those defined with the triplet  $(\alpha^n, \beta^n, \gamma^n)$ .

Repeating the stability analysis from Section 2.4 we then verify that the discrete scheme

$$\left\{ \begin{array}{l} \frac{\widehat{\mathbf{E}}_h^{n+1} - \widehat{\mathbf{E}}_h^n}{\Delta t} = cR\widehat{\mathbf{B}}_h^{n+\frac{1}{2}} - S^n(\omega_p) \frac{\widehat{\mathbf{J}}_h^{n+1} + \widehat{\mathbf{J}}_h^n}{2} \\ \frac{\widehat{\mathbf{B}}_h^{n+\frac{1}{2}} - \widehat{\mathbf{B}}_h^{n-\frac{1}{2}}}{\Delta t} = -cR^t \widehat{\mathbf{E}}_h^n \\ \frac{\widehat{\mathbf{J}}_h^{n+1} - \widehat{\mathbf{J}}_h^n}{\Delta t} = S^n(\omega_p) \frac{\widehat{\mathbf{E}}_h^{n+1} + \widehat{\mathbf{E}}_h^n}{2} + S^n(\omega_c) \mathbf{b} \wedge_h^n \frac{\widehat{\mathbf{J}}_h^{n+1} + \widehat{\mathbf{J}}_h^n}{2} \end{array} \right. \quad (\text{B.2})$$

(which can be made explicit with the procedure outlined in Section 3.3) allows to compute  $(\widehat{\mathbf{E}}_h^n, \widehat{\mathbf{B}}_h^{n-\frac{1}{2}}, \widehat{\mathbf{J}}_h^n) \mapsto (\widehat{\mathbf{E}}_h^{n+1}, \widehat{\mathbf{B}}_h^{n+\frac{1}{2}}, \widehat{\mathbf{J}}_h^{n+1})$  in such a way that the pseudo-energy (31)

$$\mathcal{E}_h^n = \|\widehat{\mathbf{E}}_h^n\|_h^2 + \|\widehat{\mathbf{B}}_h^{n-\frac{1}{2}}\|_h^2 + \|\widehat{\mathbf{J}}_h^n\|_h^2 - \Delta t \langle \widehat{\mathbf{E}}_h^n, cR\widehat{\mathbf{B}}_h^{n-\frac{1}{2}} \rangle_h \quad (\text{B.3})$$

is preserved. In particular, we see that the use of a time-dependent cluster pattern (B.1) does not change anything on the symetrized formulation. However it affects the scheme formulated in the original variables, due to the fact that being pattern-dependent the change of variables now becomes also time-dependent, namely

$$\mathbf{J}_h^n = c\varepsilon_0 S^n(\omega_p) \widehat{\mathbf{J}}_h^n. \quad (\text{B.4})$$

The first effect is that the discrete pseudo-energy, seen as a functional of the field  $\mathbf{J}_h^n$ , changes with time. The second effect is a slight modification of the numerical scheme (32). Indeed applying (B.4) we find that (32) reformulated in the original variables reads

$$\left\{ \begin{array}{l} \varepsilon_0 \frac{\mathbf{E}_h^{n+1} - \mathbf{E}_h^n}{\Delta t} = R\mathbf{H}_h^{n+\frac{1}{2}} - \frac{\tilde{\mathbf{J}}_h^{n+1} + \mathbf{J}_h^n}{2}, \\ \mu_0 \frac{\mathbf{H}_h^{n+\frac{3}{2}} - \mathbf{H}_h^{n+\frac{1}{2}}}{\Delta t} = -R^t \mathbf{E}_h^{n+1}, \\ \frac{\tilde{\mathbf{J}}_h^{n+1} - \mathbf{J}_h^n}{\Delta t} = \varepsilon_0 S^n(\omega_p^2) \frac{\mathbf{E}_h^{n+1} + \mathbf{E}_h^n}{2} + S^n(\omega_c) \mathbf{b} \wedge_h \left( \frac{\tilde{\mathbf{J}}_h^{n+1} + \mathbf{J}_h^n}{2} \right), \\ \mathbf{J}_h^{n+1} = S^{n+1}(\omega_p) S^n(\omega_p^{-1}) \tilde{\mathbf{J}}_h^{n+1}. \end{array} \right.$$

(Although it has not been used here, we may note that  $S^n(u)$  and  $S^{n+1}(v)$  commute as diagonal operators.) Clearly the explicit version of this scheme is obtained by a straightforward application of Proposition 3.4: it reads

$$\left\{ \begin{array}{l} \mathbf{H}_h^{n+\frac{1}{2}} = \mathbf{H}_h^{n-\frac{1}{2}} - \frac{\Delta t}{\mu_0} R^t \mathbf{E}_h^n \\ \tilde{\mathbf{J}}_h^{n+1} = \left( I + S^n\left(\frac{\theta}{1+\theta^2}\right) (\mathbf{b} \wedge_h) + S^n\left(\frac{\theta^2}{1+\theta^2}\right) (\mathbf{b} \wedge_h)^2 \right) \mathbf{W}_h^n \\ \mathbf{E}_h^{n+1} = \mathbf{E}_h^n + \frac{\Delta t}{\varepsilon_0} \left( R\mathbf{H}_h^{n+\frac{1}{2}} - \frac{1}{2} (\tilde{\mathbf{J}}_h^{n+1} + \mathbf{J}_h^n) \right) \\ \mathbf{J}_h^{n+1} = S^{n+1}(\omega_p) S^n(\omega_p^{-1}) \tilde{\mathbf{J}}_h^{n+1} \end{array} \right.$$

with auxiliary fields given by  $\theta = \frac{\Delta t \omega_c}{2\eta}$ ,  $\eta = 1 + \left(\frac{\Delta t \omega_p}{2}\right)^2$  and

$$\begin{aligned} \mathbf{W}_h^n &= S^n\left(\frac{1}{\eta}\right) \left( \Delta t \varepsilon_0 S^n(\omega_p^2) (\mathbf{E}_h^n + \frac{\Delta t}{2\varepsilon_0} R\mathbf{H}_h^{n+\frac{1}{2}}) \right. \\ &\quad \left. + \left( I - \frac{\Delta t^2}{4} S^n(\omega_p^2) + \frac{\Delta t}{2} S^n(\omega_c) \mathbf{b} \wedge_h \right) \mathbf{J}_h^n \right). \end{aligned}$$

## Acknowledgments

The authors acknowledge the support of ANR under contract ANR-12-BS01-0006-01. Moreover this work was carried out within the framework of the European Fusion Development Agreement and the French Research Federation for Fusion Studies. It is supported by the European

Communities under the contract of Association between Euratom and CEA, and has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement number 633053. IST activities also received financial support from "Fundação para a Ciência e Tecnologia" through project Pest-OE/SADG/LA0010/2013. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

- [1] M.U. Böhner, Simulation of Microwave Propagation in a Fusion Plasma, Master Thesis, Institut für Plasmaforschung University of Stuttgart, January 2011.
- [2] M. Branbilla, Kinetic Theory of Plasma Waves-Homogeneous Plasmas, Clarendon Press, International Series of Monographs on Physics, 1998.
- [3] F.F. Chen and R.B. White, Amplification and Absorption of Electromagnetic Waves in Overdense Plasmas, *Plasma Phys.* **16** (1974) 565-587; anthologized in *Laser Interaction with Matter*, Series of Selected Papers in Physics, ed. by C. Yamanaka, Phys. Soc. Japan, 1984
- [4] S.J. Cooke, M. Botton, T.M. Antonsen and B. Levush, A leapfrog formulation of the 3D ADI-FDTD algorithm, *Int. J. Numer. Model.* **22** (2009) 187-200.
- [5] F. da Silva, S. Heuraux, and M. Manso, Studies on O-mode reflectometry spectra simulations with velocity shear layer, *Nucl. Fusion*, **46**(9), (2006) S816-S823.
- [6] F. da Silva, S. Heuraux, E. Gusakov and A. Popov, A numerical study of forward- and back-scattering signatures on Doppler reflectometry signals, *IEEE Trans. Plasma Sci* **38** (2010) 2144-48.
- [7] F. da Silva, S. Heuraux and M. Manso, Developments on reflectometry simulations for fusion plasmas : applications to ITER position reflectometry, *J. Plasma Physics* **72** (2006) 1205.
- [8] F. da Silva, S. Heuraux, T. Ribeiro and B. Scott, Development of a 2D full-wave JE-FDTD Maxwell X-mode code for reflectometry simulation, IRW9, Lisbon, (2009).
- [9] S. Del Pino, B. Després, P. Havé, H. Jourdain and P.F. Piserchia, 3D finite volume simulation of acoustic waves in the earth atmosphere. *Comput. & Fluids*, **38**(4):765-777, 2009.
- [10] B. Després, L.M. Imbert-Gérard and R. Weder, Hybrid resonance of Maxwell's equations in slab geometry, *Journal de Mathématiques Pures et Appliquées*, Volume 101, Issue 5, May 2014, Pages 623-659.
- [11] S. Desroziers, F. Nataf and R. Sentis, Simulation of laser propagation in a plasma with a frequency wave equation *Journal of Computational Physics*, Volume 227, Issue 4, 1 February 2008, Pages 2610-2625.
- [12] R.J. Dumont, C.K. Phillips and D.N. Smithe, Effects of non-Maxwellian species on ion cyclotron waves propagation and absorption in magnetically confined plasmas, *Phys. of Plasmas* **12** (2005) 042508.
- [13] D.Y. Heh and E.L. Tan, FDTD modeling for dispersive media using matrix exponential method, *IEEE Microw. Wirel. Compon. Lett.* **19** (2009) 53-55.
- [14] S. Huang and F. Li, FDTD implementation for magnetoplasma medium using exponential time differencing, *IEEE Microw. Wirel. Compon. Lett.* **15** (2005) 183-185.
- [15] Joo Hwa Lee and Dikshitulu K. Kalluri, Three-Dimensional FDTD Simulation of Electromagnetic Wave Transformation in a Dynamic Inhomogeneous Magnetized Plasma, *IEEE antennas and wireless propagation letters*, **47** (1999) 1146-1151.
- [16] G.V. Jandieri, A. Ishimaru, N.F. Mchedlishvili and I.G. Takidze, Spatial power spectrum of multiple scattered ordinary and extraordinary waves in magnetized plasma with electron density fluctuations, *Progress In Electromagnetics Research M*, **25**, (2012) 87-100.
- [17] A. Kohn, J. Jacquot, M.W. Bongard, S. Gallian, E.T. Hinson and F.A. Volpe, Full-wave modeling of the O-X mode conversion in the Pegasus Toroidal Experiment, arXiv:1104.0743 [physics.plasm-ph](2011).
- [18] D.Liang and Q. Yuan, The spatial fourth-order energy-conserved S-FDTD scheme for Maxwell's equations, *J. Computational Physics* **243** (2013) 344-364.
- [19] S. Liu and Sh. Liu, Runge-Kutta Exponential Time Differencing FDTD Method for Anisotropic Magnetized Plasma, *IEEE antennas and wireless propagation letters*, **7** (2008) 306-309.
- [20] Y. Peysson, J. Decker, L. Morini and S. Coda, RF current drive and plasma fluctuations, *Plasma Phys. Control. Fusion* **53** (2011) 124028.
- [21] D.N. Smithe, Finite-difference time-domain simulation of fusion plasmas at radiofrequency time scales, *Phys. of Plasmas*, **14** (2007) 056104.
- [22] T.H. Stix, The theory of plasma waves, Mac Graw Hill Advanced physics monograph series, 1962.
- [23] G. Swanson, Plasma waves, IOP series in plasma physics, 2003.
- [24] E.V. Sysoeva, E.Z. Gusakov, F. da Silva and S. Heuraux, Microwave beam broadening in turbulent magnetized plasmas: comparison theory-2D full-wave computations, submitted to *Plas. Phys. Control. Fusion*.
- [25] E.L. Tan, Fundamental schemes for efficient unconditionally stable implicit finite-difference time-domain methods, *IEEE Trans. Antenn. Propag.* **56** (2008) 170-177 .
- [26] I. Terrasse and T. Abboud, Modélisation des phénomènes de propagation d'ondes, Master lecture notes of ENSTA

- high school, <http://catalogue.polytechnique.fr/Files/Poly%20MAP%20558.pdf> (2007, in french).
- [27] C. Tsironis, A.G. Peeters, H. Isliker, D. Strintzi, I. Chatziantonaki and L. Vlahos, Electron-cyclotron wave scattering by edge density fluctuations in ITER, *Phys. of Plasmas* **16**, (2009) 112510.
  - [28] L. Xu and N. Yuan, FDTD Formulations for Scattering From 3-D Anisotropic Magnetized Plasma Objects, *IEEE antennas and wireless propagation letters*, **5** (2006) 335-338.
  - [29] K.S. Yee, Numerical Solution of Initial Boundary Value Problems Involving Maxwell's Equations in Isotropic Media, *IEEE*, **14** (1966) 802-808.
  - [30] Y. Yu and J. Simpson, An EJ collocated 3-D FDTD model of electromagnetic wave propagation in magnetized cold plasma, *IEEE Trans. Antenn. Propag.* **58** (2010) 469-478.
  - [31] F. Zheng, Z. Chen and J. Zhang, Toward the development of a three dimensional unconditionally stable finite-different time-domain method, *IEEE Trans. Microw. Theory Tech.* **48** (2000) 1550-1558 .